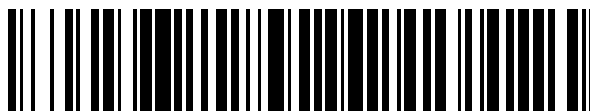


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 784 233**

51 Int. Cl.:

G05B 13/04 (2006.01)

G05B 17/02 (2006.01)

G06Q 50/06 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.09.2015 E 15306546 (1)**

97 Fecha y número de publicación de la concesión europea: **08.01.2020 EP 3112960**

54 Título: **Procedimiento combinado para detectar anomalías en un sistema de distribución de agua**

30 Prioridad:

29.06.2015 EP 15306029

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

23.09.2020

73 Titular/es:

**SUEZ GROUPE (100.0%)
16 Place de l'Iris Tour CB21
92040 Paris la Défense CEDEX, FR**

72 Inventor/es:

**CUSSONNEAU, GUILLAUME;
JARRIGE, PIERRE-ANTOINE;
DEMBELE, ABEL y
CAMPAN, FRANCIS**

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 784 233 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Procedimiento combinado para detectar anomalías en un sistema de distribución de agua

Campo de la invención

5 La presente invención se refiere a la detección de anomalías en un sistema de distribución de agua. Más específicamente, se refiere a la detección de anomalías con el uso combinado de estadísticas y un modelo hidráulico del sistema de distribución de agua.

Antecedentes**Técnica anterior**

10 Un sistema para distribuir agua potable está hecho principalmente de tuberías entre un cabezal hidráulico y los consumidores, junto con dispositivos de control tales como válvulas y bombas. Dicho sistema puede estar sujeto a numerosas anomalías. Las anomalías pertenecen a una pluralidad de tipos. Las anomalías hidráulicas comprenden fugas, variación anormal de la presión, caída rápida del nivel del agua de un tanque de almacenamiento, balance de materia de almacenamiento incoherente. Las anomalías de operación definen un elemento en el sistema que está en un estado incorrecto, por ejemplo, una válvula en un estado de apertura diferente respecto de la almacenada en el sistema de información. Estas anomalías, especialmente las fugas, pueden reducir enormemente el rendimiento del sistema de distribución de agua. Por ejemplo, las fugas en las tuberías son la causa de la pérdida de una parte significativa del agua entre el cabezal hidráulico y los consumidores, y pueden causar daños estructurales. La detección y corrección de anomalías en un sistema de distribución de agua es, por lo tanto, una preocupación permanente de los operadores de dichos sistemas para mitigar el coste económico de la pérdida y daños del agua. Además, la detección de fugas en un sistema de distribución de agua es un objetivo fundamental para limitar el consumo global de agua y los residuos, lo cual es de particular interés en las regiones sujetas a escasez de agua y con el objetivo de promover el desarrollo sostenible.

15 La detección de fugas en un sistema de distribución de agua se realiza históricamente mediante inspección humana. La inspección humana generalmente consiste en enviar operadores humanos para inspeccionar las tuberías del sistema e identificar fugas y otras anomalías. Esta detección se puede facilitar, por ejemplo, con sensores de audio, que detectan ruido debido a una fuga. Sin embargo, el tamaño típicamente grande de los sistemas de distribución de agua hace que la detección humana de fugas y anomalías sea muy difícil. Por ejemplo, el sistema de distribución de agua de una gran ciudad comprende miles de kilómetros de tuberías. Por lo tanto, es imposible inspeccionar todas las tuberías con frecuencia a un coste razonable.

20 El uso de uno o más sensores contribuye a una solución al problema de detectar anomalías en un sistema de distribución de agua. Los sensores pueden usarse principalmente para detectar automáticamente cambios anormales en el comportamiento del sistema, por lo tanto, se pueden enviar operadores humanos a los nodos/arcs del sistema en los que se detectó un comportamiento anormal. Sin embargo, los procedimientos de detección de anomalías basados en sensores también tienen desventajas. Pueden generar una gran cantidad de falsos positivos (alarmas para eventos que no son anomalías reales) y dar lugar a numerosas intervenciones humanas inútiles y costosas. También pueden detectar anomalías correctamente pero con localización imprecisa, debido a la gran escala de un sistema típico de distribución de agua. De otra manera, implementar una alta densidad de sensores en el sistema podría ayudar a localizar anomalías con mayor precisión, pero es demasiado costoso.

25 El documento US 2014/0052421 desvela un procedimiento para modelar un sistema de distribución de agua. Este procedimiento consiste en simular la red hidráulica utilizando estimaciones de la demanda de agua. El procedimiento comprende una etapa de calibración, que consiste en modificar las demandas de agua y simular nuevamente el sistema hidráulico, con el fin de minimizar las diferencias entre las predicciones y observaciones de los parámetros hidráulicos, como las presiones y los caudales. El procedimiento que desvela el documento US 2014/0052421 permite calibrar el modelado de una red de distribución de agua, pero no permite detectar anomalías en la red de distribución de agua.

30 La solicitud de patente europea n.º EP 15306029.8, presentada por el mismo solicitante que el solicitante de la presente solicitud, desvela procedimientos para detectar anomalías en un sistema de distribución de agua. Un procedimiento desvelado en esta solicitud anterior utiliza un conjunto de valores predichos de parámetros de entidades del sistema de distribución de agua, modelado como una red, por ejemplo, la velocidad del agua y la presión del agua en diferentes momentos, de acuerdo con un modelo hidráulico del sistema de distribución de agua y los valores de los parámetros de control de la red, por ejemplo la longitud, rugosidad de las tuberías, una predicción del consumo de los usuarios de la red, etc. Dicho procedimiento consiste en realizar iteraciones sucesivas del cálculo de los valores predichos, comparar estos valores con observaciones de los sensores y, de acuerdo con las diferencias entre predicciones y observaciones, ajustar los valores de los parámetros, usando luego el procedimiento para detectar una anomalía de acuerdo con los valores de los parámetros. Este procedimiento permite una determinación precisa de los valores de los parámetros de control de la red que mejor coinciden con las observaciones de los sensores. Algunos valores de los parámetros de control son representativos de anomalías en la red. Por ejemplo, la predicción de un consumo excesivo de agua puede ser representativa de una fuga en el sistema de distribución de agua. Los algoritmos

de aprendizaje automático, capacitados en casos ejemplares de anomalías en una red de agua, son capaces de identificar una anomalía de acuerdo con los valores de los parámetros de control que se obtienen después de varios casos de cálculo de los valores predichos, a través del procedimiento iterativo.

5 Este procedimiento tiene muchas ventajas. Proporciona un procedimiento automático, rápido y fiable para detectar anomalías en un sistema de distribución de agua, lo que permite corregir anomalías en el sistema mucho más rápidamente que con el uso de inspecciones humanas tradicionales. Sobre iteraciones sucesivas, las modificaciones para controlar los parámetros se identifican cada vez con mayor precisión. Este procedimiento es eficiente para identificar y localizar anomalías en un sistema de distribución de agua después de sucesivas iteraciones. Además, el uso del modelo hidráulico del sistema de distribución de agua permite al operador de una red localizar una anomalía
10 incluso en una entidad de la red que no está equipada con sensores.

Sin embargo, la gran cantidad de parámetros que tienen un impacto en el modelo hidráulico y la gran cantidad de iteraciones necesarias para obtener una solución satisfactoria pueden hacer que el procedimiento sea complejo en grandes sistemas de distribución de agua. Esto mitiga la capacidad del operador de un sistema de distribución de agua para identificar y corregir anomalías en el sistema de distribución de agua rápidamente.

15 Por lo tanto, existe la necesidad de un procedimiento para disminuir el tiempo de detección notificado al operador y mejorar la relevancia de los resultados en términos de identificación y localización. Por ejemplo, el tiempo de detección se reduce con la optimización del número de iteraciones de cálculo de los parámetros de control de una red de sistema de distribución de agua que son necesarios para lograr una detección fiable de anomalías en el sistema de distribución de agua.

20 **Sumario de la invención**

A tal efecto, la invención desvela un procedimiento para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho procedimiento: parametrizar un modelo hidráulico del sistema de distribución de agua con un conjunto de valores de variables de control que caracterizan la red y su salida en los nodos; utilizar sensores en la red para adquirir observaciones de un subconjunto de variables de estado en primeras
25 referencias temporales; identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones; cambiar el conjunto de valores de las variables de control utilizando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado; realizar una clasificación de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control.

30 Ventajosamente, el procedimiento comprende además identificar al menos una variable de control para cambiar en función de al menos dichas observaciones.

Esto permite una reducción adicional de la carga de cálculo del procedimiento, reduciendo el número de variables de control que se han de modificar.

35 Ventajosamente, identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende: usar el modelo hidráulico para calcular los valores predichos de un conjunto de variables de estado que caracterizan al menos una velocidad y presión del agua en los nodos, asociándose dichos valores predichos a segundas referencias temporales; calcular los valores residuales del subconjunto de las variables de estado como una diferencia entre los valores predichos y las observaciones en las segundas referencias temporales; realizar un análisis estadístico de valores residuales en una entidad de la red en
40 una selección de las referencias temporales; clasificar las entidades de la red en función de las reglas aplicadas a la salida del análisis estadístico.

Esto permite una detección previa eficiente de al menos una entidad objetivo, utilizando valores históricos.

Ventajosamente, dicha al menos una entidad objetivo se identifica clasificándose en un estado anormal en función de la comparación de valores residuales y un umbral predefinido.

45 Esto permite una selección eficiente de al menos una entidad objetivo, centrándose en las entidades que tienen mayores posibilidades de estar en un estado anormal.

Ventajosamente, identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende: realizar una serie de pruebas en las observaciones utilizando una combinación de reglas; clasificar las entidades de la red en función de la combinación de reglas sobre
50 observaciones.

Esto permite una clasificación simple pero robusta de entidades sin la necesidad de utilizar un modelo hidráulico.

Ventajosamente, dicha al menos una entidad objetivo se identifica clasificándose en un estado anormal en función de la combinación de reglas sobre observaciones.

Esto permite una identificación simple pero robusta de al menos una entidad objetivo.

- Ventajosamente, identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende: realizar una serie de pruebas sobre desviaciones estadísticas de observaciones utilizando una combinación de reglas; clasificar las entidades de la red en función de la combinación de reglas sobre observaciones.
- 5 Esto permite una clasificación simple pero robusta de entidades sin la necesidad de utilizar un modelo hidráulico.
- Ventajosamente, dicha al menos una entidad objetivo se identifica al clasificarse en un estado anormal en función de la combinación de reglas sobre desviaciones estadísticas de observaciones.
- Esto permite una identificación simple pero robusta de al menos una entidad objetivo.
- 10 Ventajosamente, identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende identificar un arco de la red en el que el valor de una observación de la velocidad del agua excede un umbral en una entidad objetivo.
- Esto permite una identificación muy simple y directa de al menos una entidad objetivo.
- Ventajosamente, el conjunto de valores de las variables de control se realiza para al menos un tipo de evento, y las variables de control que se han de modificar se basan en dicho al menos un tipo de evento.
- 15 Esto permite mejorar la robustez de la detección y disminuir la carga de cálculo del procedimiento. De hecho, se modifica un número reducido de variables de control, que tienen una mayor probabilidad de vincularse a un evento producido en la red.
- Ventajosamente, se prueba una pluralidad de tipos de eventos y se realiza una instancia de cambio del conjunto de valores de variables de control para cada tipo de evento.
- 20 Esto permite probar al mismo tiempo una pluralidad de diferentes tipos de eventos.
- Ventajosamente, al menos cambiar el conjunto de valores de las variables de control se realiza para al menos un tipo de evento elegido mediante la clasificación de las entidades de la red en función de las reglas aplicadas a la salida del análisis estadístico.
- 25 Esto permite reducir la carga de cálculo de cambiar el conjunto de valores de las variables de control, determinando más específicamente las variables de control que se han de cambiar, de acuerdo con el análisis estadístico.
- Ventajosamente, el procedimiento comprende: detectar una anomalía en función del resultado de identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones; si no se detecta ninguna anomalía, enriquecer una base de datos de estados normales de las entidades; si se detecta una anomalía: cambiar el conjunto de valores de las variables de control utilizando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de los valores residuales de las variables de estado; realizar la segunda clasificación de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control.
- 30 Esto permite reducir la carga de cálculo del procedimiento, iniciando el ajuste iterativo de las variables de control solo si se detecta una anomalía en función de las observaciones, y mejorar las ejecuciones posteriores del procedimiento al enriquecer una base de datos de estados normales.
- 35 Ventajosamente, el cambio del conjunto de valores de las variables de control mediante un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado comprende: A) cambiar el conjunto de valores de las variables de control; B) usar el modelo hidráulico para calcular los valores predichos de un conjunto de variables de estado que caracterizan al menos la velocidad y la presión del agua en los nodos en las referencias temporales; C) calcular los valores residuales del conjunto de variables de estado como una diferencia entre los valores predichos y las observaciones en las referencias temporales; D) si dichas diferencias satisfacen un criterio de interrupción, pasar a la etapa F); E) si no, cambiar el conjunto de valores de las variables de control (350) y volver a la etapa B); F) si dichas diferencias no satisfacen un criterio de refinamiento (360), seleccionar un subconjunto de la red (370) en que calcular los valores predichos, volviendo a la etapa B).
- 40 Esto permite un ajuste progresivo de los valores de las variables de control hasta que sus valores sean lo suficientemente fiables como para detectar eventos en la red.
- Ventajosamente, el criterio de refinamiento comprende el cálculo de los valores de uno de un mínimo cuadrado y una función objetivo bayesiana, y la selección y modificación de las variables de control se determinan mediante un algoritmo de Levenberg-Marquardt.
- 45 El uso de una función objetivo bayesiana permite calibrar la fiabilidad relativa de las mediciones y los valores modificados de las variables de control y estado. El uso del algoritmo de Levenberg-Marquardt permite tener buenas cualidades de convergencia y estabilidad.
- 50

Ventajosamente, la realización de una clasificación de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control se realiza mediante un algoritmo de aprendizaje automático previamente capacitado.

Esto permite tener reglas de clasificación optimizadas para un sistema de distribución de agua dado, cuya fiabilidad aumenta con el tiempo.

- 5 Ventajosamente, las variables de control comprenden variables escalares que caracterizan la topología y la topografía de la red, y variables basadas en el tiempo que caracterizan las entradas y salidas de la red que tienen al menos un valor en cada referencia temporal.

Esto permite aumentar el número de posibles eventos detectados por el procedimiento.

- 10 Ventajosamente, cambiar el conjunto de valores de las variables de control comprende modificar al menos uno de: los valores de un subconjunto de las variables de control basadas en el tiempo; los valores de un subconjunto de las variables de control escalar calculados durante una fase de modelado de la red.

Esto permite aprovechar al máximo la disponibilidad de variables de control escalares y basadas en el tiempo.

Ventajosamente, modificar los valores de un subconjunto de las variables de control basadas en el tiempo comprende modificar los valores de las variables de control basadas en el tiempo representativas del consumo de agua.

- 15 Esto permite detectar eventos que se detectan por un aumento anormal del consumo de agua, por ejemplo fugas.

Ventajosamente, las variables de estado caracterizan aún más la presión.

Esto permite tener un modelo más completo y eficiente de la evolución temporal de los parámetros físicos del sistema de distribución de agua.

- 20 La invención también desvela un sistema para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho sistema: sensores de al menos velocidad y presión del agua en un subconjunto de nodos de la red; un dispositivo informático que comprende un procesador; enlaces de comunicación entre sensores y el dispositivo informático; un medio de almacenamiento; en el que el dispositivo informático está configurado para: recuperar un conjunto inicial de valores de variables de control que caracterizan la red y su salida en los nodos de los medios de almacenamiento y lo utilizan para parametrizar un modelo hidráulico del sistema de distribución de agua; utilizar enlaces de comunicación entre sensores en la red para adquirir observaciones de un subconjunto de las variables de estado, teniendo dichas observaciones referencias temporales; ejecutar uno de los procedimientos desvelados anteriormente.
- 25

- 30 La invención también desvela un producto de programa informático, almacenado en un medio no transitorio legible por ordenador, para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho producto de programa informático instrucciones de código para ejecutar uno de los procedimientos desvelados anteriormente.

La invención aumenta la robustez y reduce el coste de cálculo de un procedimiento de detección de anomalías en un sistema de distribución de agua basado en un ajuste paso a paso de los parámetros de control.

- 35 La invención permite una detección precisa de anomalías en cualquier entidad de un sistema de distribución de agua, incluso si dicha entidad no está equipada con sensores.

Los procedimientos de la invención son capaces de detectar varios tipos de anomalías en un sistema de distribución de agua.

La precisión de la detección de anomalías de un procedimiento de la invención aumenta con el número de usos del procedimiento.

40 **Breve descripción de los dibujos**

La invención se entenderá mejor y sus diversas características y ventajas surgirán de la siguiente descripción de una serie de realizaciones ejemplares y sus figuras adjuntas, en las que:

- la figura 1 muestra un ejemplo de un sistema de distribución de agua organizado en red en la técnica anterior;
 - la figura 2 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, usando un ajuste iterativo paso a paso de propiedades y variables de control de un modelo hidráulico combinado con un aprendizaje automático para identificar la anomalía en la técnica anterior;
 - la figura 3 muestra un ejemplo de un procedimiento para identificar parámetros para un evento objetivo en la técnica anterior;
 - la figura 4 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua
- 45
- 50

- en una serie de realizaciones de la invención;
- la figura 5 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, en la que la ejecución del ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico combinado con un aprendizaje automático para identificar la anomalía depende de la salida de una clasificación de anomalías utilizando un cálculo de los valores residuales de las variables de estado;
 - la figura 6 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, en la que la ejecución del ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico combinado con un aprendizaje automático para identificar la anomalía depende de la salida de una clasificación de anomalías utilizando una combinación de reglas aplicadas a las variables de estado;
 - la figura 7 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, utilizando el aprendizaje estadístico para detectar y caracterizar previamente una anomalía.
 - la figura 8 muestra un ejemplo de una presentación de la localización de anomalías a un operador.

Descripción detallada de la invención

En esta memoria descriptiva, la invención se describirá mediante ejemplos relacionados con la detección de fugas y la detección de concentraciones anormales de cloro. Sin embargo, la invención no se limita a estos ejemplos y puede aplicarse a la detección de cualquier anomalía en un sistema de distribución de agua.

La figura 1 muestra un ejemplo de un sistema de distribución de agua organizado en una red de acuerdo con la técnica anterior.

La red 100 mostrada en la figura 1 está organizada como una red y comprende una pluralidad de nodos 110, 111, 112, 113, 114, 115, 116 y 117, así como una pluralidad de arcos 120, 121, 122, 123, 124, 125, 126, 127 y 128. Los nodos suelen representar las conexiones con fuentes de agua o depósitos de agua, por ejemplo el depósito 130 en el nodo 116, las conexiones con el usuario del sistema de distribución de agua, por ejemplo, el consumo 131 en el nodo 113 y las conexiones entre los arcos. Los arcos típicamente representan tuberías entre los nodos. La red puede equiparse con equipos como válvulas y bombas. Por ejemplo, una bomba 132 está presente en el arco 120. Más generalmente, un nodo puede ser una unión entre dos o tres tuberías, un punto en el que se encuentran las entradas o salidas de la red, por ejemplo, un punto en el que un usuario consume agua, o un punto en el que se inyecta agua en la red. Un nodo también puede representar una subred, por ejemplo, un vecindario agrupado en un solo nodo.

Los parámetros físicos relacionados con el agua comprenden notablemente a modo de ejemplo la velocidad, presión, caudal, nivel de almacenamiento (depósito y tanque), temperatura, etc. Se pueden agregar parámetros adicionales de acuerdo con las necesidades de detección y/o la evolución de los sensores. La evolución de estos parámetros a lo largo del tiempo depende de las características del sistema de distribución de agua, las entradas y salidas en los nodos y el estado de cualquier equipo en el sistema.

Los atributos del sistema de distribución de agua comprenden principalmente:

- la topología de la red (es decir, el gráfico con arcos y nodos);
- la topografía de la red (es decir, la elevación de los nodos);
- las propiedades de las tuberías, como la longitud, diámetro, material, rugosidad, singularidades, coeficientes de pérdida menores, etc.);
- las propiedades de cada equipo de la red (por ejemplo, características de la bomba, diámetro de la válvula, coeficiente de fricción, consigna de funcionamiento, etc.).

El modelo hidráulico de una red de agua potable comprende este conjunto de información.

Las entradas y salidas en los nodos están definidas por los consumos o inyecciones de agua en cada nodo del gráfico. Típicamente, representan los consumos individuales de los usuarios del sistema de distribución de agua y la inyección de agua desde la entrada de agua al sistema.

El estado del equipo, como válvulas o bombas, comprende los estados operativos y los puntos de consigna de estos dispositivos.

En esta descripción, las características del sistema de distribución de agua, las entradas y salidas en los nodos y estados de los dispositivos se denominarán "variables de control" del sistema de distribución de agua, mientras que los parámetros físicos en los nodos y arcos se denominarán "variables de estado" del sistema de distribución de agua. Las variables de control comprenden, por ejemplo, el consumo de agua en los nodos, expresado, por ejemplo, en m³/d, rugosidad y fricción de las tuberías, así como el estado de las bombas o accionadores, etc. Las variables de estado comprenden velocidad y presión en los diferentes arcos y nodos. De acuerdo con la aplicación objetivo, las variables de estado también pueden comprender niveles de agua en tanques y depósitos de agua, concentración de cloro residual (RCC), etc. Los valores de las variables de estado varían con el tiempo. Los valores de algunas variables de control, por ejemplo, aquellos relacionados con la topología o la topografía de la red, permanecen constantes,

mientras que los valores de algunas variables de control, por ejemplo, aquellos relacionados con las entradas y salidas de la red y con los estados de los dispositivos, varían con el tiempo. Mientras tanto, una entidad se refiere a toda la red, o un subconjunto de una red compuesta de nodos, arcos y equipos, un solo nodo, un solo arco o un equipo de la red que se caracteriza por características que tienen un impacto en el comportamiento del sistema de distribución de agua.

Una descripción del valor inicial de las variables de estado y una predicción (serie temporal) de los valores de las variables de control, principalmente los consumos en los nodos de la red, permiten una predicción de los valores de las variables de estado a lo largo del tiempo. Esta predicción generalmente se realiza calculando valores sucesivos de las variables de estado con una etapa de tiempo fija o variable. Esta predicción se llama resolución del problema directo.

Algunas entidades de la red están equipadas con sensores que miden parámetros físicos, típicamente velocidad o presión. Estos sensores permiten obtener series temporales de variables de estado.

Se conocen procedimientos para calcular los valores de las variables de control que definen la bondad de ajuste del sistema además de las observaciones de las variables de estado. Estos procedimientos generalmente se denominan resolución inversa de problemas y aparecen desvelados principalmente en Piller, O. (1995): Modeling the behavior of a network -Hydraulic analysis and sampling procedures for parameter estimation. Applied Mathematics thesis from the University of Bordeaux (PRES), 288 páginas, Talence, Francia, presentado el 3 de febrero de 1995 y Piller, O., Gilbert, D. y Van Zyl, J. E. (2010): Dual Calibration for Coupled Flow and Transport Models of Water Distribution Systems. Water Distribution Systems Analysis WDSA 2010, ASCE, Tucson, 722-731, diciembre de 2010, US 8635051. Resolver un problema inverso generalmente consiste en iteraciones sucesivas de un bucle que comprende modificar valores de variables de control, predecir valores de variables de estado de acuerdo con las variables de control, para minimizar una función matemática que sea representativa de las diferencias entre predicción y observaciones de variables de estado. Estas técnicas permiten determinar, por ejemplo, las características de una entidad de la red (por ejemplo, una pérdida de carga en una válvula, una rugosidad de una tubería, etc.) o el consumo de agua que mejor se ajusta a las observaciones de las variables de estado.

En esta solicitud, una entidad es un término general para designar un nodo, un arco, equipo, un subconjunto de la red o de toda la red. De este modo, clasificar una entidad en un estado anormal puede consistir en identificar una anomalía de un nodo, un arco, un equipo, un subconjunto de la red, etc.

La figura 2 muestra un ejemplo de un procedimiento para detectar anomalías en un sistema de distribución de agua en la técnica anterior, utilizando un ajuste iterativo paso a paso de propiedades y variables de control de un modelo hidráulico y aprendizaje automático para identificar la anomalía.

El procedimiento 200 comprende una primera etapa 220 para definir parámetros de un modelo hidráulico 211 usando un conjunto de variables de control 210. La naturaleza del modelo hidráulico y las variables de control se han presentado en relación con la figura 1. El modelo hidráulico puede configurarse principalmente con variables de control que caracterizan la estructura de la red y variables de control que caracterizan una predicción de las entradas y salidas de la red en los nodos durante un conjunto de referencias temporales, principalmente una predicción del consumo de agua en la red.

En una realización preferente, los valores de las variables de control relacionadas con la estructura de la red (por ejemplo, el diámetro y la rugosidad de las tuberías, características de los equipos, etc.) han sido calibrados después de la creación del modelo hidráulico. De hecho, los valores teóricos definidos durante el modelado de la red no siempre coinciden con los valores reales. Para obtener valores más precisos de las variables de control relacionadas con la estructura de la red, una etapa de calibración puede consistir en ajustar los valores de estas variables de control para proporcionar la mejor predicción del comportamiento de la red. Típicamente, este ajuste puede consistir en:

- realizar observaciones de las entradas, salidas y un subconjunto de las variables de estado de la red durante una duración determinada;
- configurar el modelo hidráulico de la red con valores de las variables de control relacionadas con las entradas, salidas y estructura de la red;
- calcular los valores predichos de las variables de estado de la red de acuerdo con el modelo hidráulico;
- calcular una diferencia entre los valores observados y predichos de las variables de estado de la red, y modificar los valores de las variables de control de la red utilizando un algoritmo de optimización para minimizar la distancia entre los valores observados y predichos de las variables de estado de la red.

Por otro lado, las variables de control relacionadas con la entrada y salida de la red se pueden obtener, por ejemplo, utilizando observaciones de las entradas y salidas pasadas. Por ejemplo, los valores de las variables de control que caracterizan el consumo de agua en diferentes nodos se pueden calcular utilizando datos medidos históricos del consumo de agua.

El procedimiento 200 comprende además una etapa 230 de usar sensores para adquirir observaciones de un subconjunto de variables de estado, teniendo dichas observaciones primeras referencias temporales. Una variable de estado es un valor temporizado de un parámetro físico en una entidad de la red. La variable de estado a menudo se

relaciona con valores físicos en un nodo de la red (por ejemplo, presión o concentración de cloro en un nodo de la red). También pueden relacionarse con un valor físico en un arco de la red, por ejemplo, la velocidad del agua en una tubería. Partes de los procedimientos de la técnica anterior y la invención se basan en una comparación de observaciones y predicciones de variables de estado. Las observaciones de las variables de estado solo son posibles si está presente un sensor adecuado. Por ejemplo, una observación de presión en un nodo de la red solo se puede realizar si un sensor de presión está presente en este nodo. Por otro lado, es posible calcular valores predichos de variables de estado en nodos u otras entidades que no están equipadas con sensores, utilizando un modelo hidráulico. Las primeras referencias temporales están asociadas a cada sensor, que es capaz de producir mediciones en diferentes tiempos/ritmos. De acuerdo con diversas realizaciones de la invención, las primeras referencias temporales de cada sensor pueden estar sincronizadas, o cada sensor puede tener sus propias referencias temporales. En ambos casos, la ejecución de las próximas etapas del procedimiento de la invención será la misma.

Tal y como se indica en relación con la figura 1, un sistema de distribución de agua generalmente está equipado con sensores. Estos sensores miden parámetros físicos como la velocidad, caudal, presión, etc. en los nodos o los arcos de la red, siendo dichos parámetros físicos en un nodo o un arco variables de estado de la red. Por razones prácticas de coste, mantenimiento y facilidad de operación, todos los nodos y arcos no están equipados con sensores, y aquellos equipados con sensores generalmente no están equipados con un sensor para cada parámetro físico. El número y la distribución de los sensores se pueden elegir de acuerdo con la definición de sectores del sistema de distribución de agua. La adquisición de observaciones se puede realizar de forma remota. Por ejemplo, los sensores que se implementan a lo largo de la red pueden enviar los valores de las variables de estado a una plataforma remota a través de diversos medios de telecomunicaciones, generalmente enlaces inalámbricos.

Los valores de las variables de estado varían con el tiempo, cada valor tiene una referencia temporal que representa el tiempo en que se midió el valor. En una realización de la invención, los valores de las variables de estado se adquieren utilizando un período de tiempo predefinido y se envían regularmente, posiblemente pero no necesariamente, usando otro período. Por ejemplo, los valores de las variables de estado para las que está presente un sensor pueden adquirirse y enviarse cada dos minutos, 5 minutos, 15 minutos o cada hora cada 24 horas. En una realización preferente de la invención, los sensores se sincronizan para adquirir mediciones simultáneamente. Esto permite que la plataforma remota adquiera observaciones de los valores del subconjunto de variables de estado para las cuales hay un sensor disponible en cada referencia temporal. Los valores se pueden enviar inmediatamente después de ser medidos. También se pueden almacenar localmente a nivel de dispositivo de transmisión, y luego enviarse a intervalos regulares, por ejemplo, enviando todos los valores que han sido capturados durante 15 minutos, 1 hora o cualquier otro intervalo.

El procedimiento 200 comprende además una etapa 240 de establecer al menos un objetivo. La etapa tiene como objetivo definir eventos específicos que pueden ser, por ejemplo, una variación anormal de la presión, la detección de una válvula en un falso estado de apertura. Estos eventos específicos serán los que se buscarán en las siguientes etapas del procedimiento, para varias entidades (toda la red, subconjunto de la red, etc.). Tal y como se explicó con referencia a la figura 3, se define un conjunto de parámetros para cada tipo de evento objetivo en la etapa 310.

El procedimiento 200 comprende además una etapa 250 de identificación de parámetros. En una serie de realizaciones de la invención, los parámetros se establecen para eventos específicos. En los casos en que se aborda una pluralidad de eventos, se ejecuta una etapa 250 de identificación para cada evento objetivo, ya sea uno tras otro, o en paralelo. A continuación, se proporciona un ejemplo de procedimiento para identificar parámetros en una instancia, con referencia a la figura 3. Existen varios procedimientos para determinar los eventos objetivo. En una serie de realizaciones de la invención, los eventos se recogen en una lista predefinida. En otras realizaciones, se selecciona un evento para una aplicación específica. Por ejemplo, solo se puede seleccionar un evento de "fuga". En otras realizaciones más de la invención, un operador selecciona los eventos que le interesan en una lista de todos los eventos posibles. En otras realizaciones más de la invención, se seleccionan los eventos que son detectables de acuerdo con los sensores presentes en la red. En otras realizaciones de la invención, no se busca ningún evento en particular. De este modo, se puede seleccionar un conjunto predefinido de parámetros.

El procedimiento comprende una etapa 260 para extraer y preprocesar un vector de características para preparar la etapa de clasificación. El procedimiento busca normalizar los datos y reducir la dimensión del problema. Los datos de entrada incluyen todos los resultados del modelo (véase la figura 3) y los datos de salida son un vector con un número reducido de dimensiones. En una realización, la reducción de la dimensión del problema se puede resolver utilizando un análisis de componentes principales, un análisis discriminante lineal o un algoritmo de los K vecinos más próximos.

El procedimiento comprende una etapa 270 para calificar el estado actual de una entidad. De acuerdo con la definición de una entidad propuesta anteriormente, la clasificación de una entidad puede aplicarse a toda la red, un solo arco, un solo nodo o un subconjunto de los nodos y arcos de la red. En una serie de realizaciones de la invención, el subconjunto de nodos y arcos de la red que forman la entidad se eligen para formar un conjunto coherente. Por ejemplo, una entidad puede ser un subconjunto de un área que entrega agua a un vecindario, con una sola entrada y un solo arco de salida para toda la entidad.

Se pueden usar varias técnicas de aprendizaje automático dependiendo de la disponibilidad y la cantidad de datos históricos etiquetados, compuestos de vectores de características. En una serie de realizaciones, estas técnicas

- incluyen detección de anomalías, agrupación y clasificación. En una serie de realizaciones, las reglas del aprendizaje automático se basan en algoritmos de aprendizaje automático 271 previamente capacitados en vectores de características históricas etiquetadas. Por ejemplo, en el caso de clases históricas etiquetadas suficientemente disponibles, los algoritmos de clasificación pueden analizar las clases históricas disponibles y los vectores de características correspondientes para construir un clasificador que determine automáticamente las reglas de clasificación de las entidades. Un nuevo vector de características se puede clasificar automáticamente en la clase actual relevante, por lo tanto, caracterizando el estado actual y brindando información valiosa a un operador. Por ejemplo, la clasificación puede basarse en máquinas de vectores de soporte, bosques aleatorios; regresión logística, etc. En una serie de realizaciones, se puede utilizar la clasificación con dos clases (estado normal/anormal). En una serie de realizaciones, las clasificaciones también pueden consistir en tipos de anomalías (presión, alto/bajo consumo y calidad del agua), etc. La clasificación también puede incluir un índice de fiabilidad, lo que indica la certeza de la clasificación de la anomalía. En una serie de realizaciones, la capacitación del algoritmo de aprendizaje automático se puede activar en cualquier momento según la base de tiempo (frecuencia definida), en criterios de rendimiento o en evolución en el conjunto de datos utilizado.
- 5 En el caso de clases históricas etiquetadas y vectores de características correspondientes que son insuficientes, se puede usar agrupación o detección de anomalías. La agrupación puede incluir técnicas como K-medias o agrupación jerárquica. La agrupación utiliza estados pasados, representados por vectores de características anteriores pero sin etiquetas conocidas. En una realización de la invención, el estado se declara como anormal si no está presente en agrupaciones más grandes (se prueba un criterio del tamaño de las agrupaciones).
- 10 El procedimiento comprende una etapa 280 para determinar si la clase calculada para una entidad corresponde a un estado anormal ("evento") de acuerdo con la salida de los algoritmos de aprendizaje automático o si uno de los estados se detecta como anomalía. Si se detecta un evento, el procedimiento pasa a la etapa 290. Si los estados están calificados como normales, el procedimiento pasa a la etapa 291.
- 15 El procedimiento comprende una etapa 290 para mostrar información estructurada al operador. Luego se puede presentar una lista de entidades priorizadas en estado anormal a un operador para lanzar operaciones de mantenimiento adecuadas a las entidades con la mayor probabilidad de anomalía. Estas entidades con estado anormal se presentan con información adicional recopilada a lo largo de las etapas anteriores (en función de la información contenida en los resultados del modelo y el vector de características). En una realización, la información contiene indicaciones sobre la cuantificación del nivel de anormalidad (por ejemplo, el valor del flujo de una fuga, el valor de la pérdida de carga), tiempo y duración de la anormalidad, ubicación conocida y propagación de la anormalidad, etc. Teniendo en cuenta la configuración del contexto, esta información se utiliza para priorizar los estados anormales y proporcionar un grado de anormalidad al usuario.
- 20 El procedimiento comprende una etapa 291 para enriquecer una base de datos histórica de vectores de características etiquetadas como "normales". Estos vectores de características se utilizan como entradas para el procedimiento de aprendizaje de algoritmos. Esto permite ventajosamente tener una base de datos más rica con el tiempo y, de este modo, una predicción más fiable de estados normales o anormales.
- 25 El procedimiento comprende una etapa 292 para enriquecer la base de datos histórica de vectores de características no etiquetadas como "normales" por el procedimiento. El procedimiento comprende entonces una etapa de validación de eventos por parte de los operadores. En una realización, los operadores pueden introducir el procedimiento, para cada entidad, si confirman el evento y su clase. Los vectores se pueden etiquetar de acuerdo con la clase a la que pertenecen (fuga, anomalía de presión, anomalía en la calidad del agua, etc.). La base de datos histórica de vectores de características se enriquece con el nuevo vector clasificado de características cualquiera que sea el estado "anormal" o "normal". Estos vectores de características se utilizan como entradas en el procedimiento de capacitación de algoritmos. Si no está etiquetado por el operador, los vectores todavía se almacenan pero sin etiqueta específica.
- 30 Luego se pueden usar en modos no supervisados o semisupervisados (agrupación y detección de anomalías).
- 35 La figura 3 muestra un ejemplo de un procedimiento para identificar parámetros para un evento objetivo en la técnica anterior.
- 40 El procedimiento 250 comprende una etapa 310 de establecer el problema de acuerdo con uno de los eventos específicos y entidades específicas elegidas en la etapa 240. La etapa tiene como objetivo definir los conjuntos de parámetros correspondientes al tipo de evento objetivo para la entidad. Se define un conjunto de parámetros por cada objetivo de evento y entidad. En una serie de realizaciones de la invención, un tipo de evento está asociado con una lista de parámetros, y el conjunto de parámetros se recupera de acuerdo con el evento que abarca. Por ejemplo, un evento puede ser un estado de apertura falsa de una válvula que se sabe que se caracteriza por valores inconsistentes de variables de control representativas de rugosidad y fricción, mientras que otro tipo de eventos pueden ser fugas que se sabe que se caracterizan por valores anormales de la demanda de agua.
- 45 El procedimiento 250 comprende además una etapa 320 de usar el modelo hidráulico para calcular valores predichos de un conjunto de variables de estado que caracterizan entidades, al menos velocidad en los arcos y presión en los nodos en segundas referencias temporales. Esta etapa consiste en calcular un valor predicho de las variables de estado de la entidad (red, sus subconjuntos, etc.), en las segundas referencias temporales. El usuario puede elegir las

segundas referencias de tiempo independientemente de las primeras referencias de tiempo. Por ejemplo, las segundas referencias temporales están separadas por un intervalo predefinido y constante (es decir, los valores predichos se calculan cada 1 minuto, 5 minutos, o cualquier otra duración) y no están necesariamente sincronizados con las primeras referencias temporales. Los procedimientos para usar un modelo hidráulico para calcular los valores predichos de las variables de estado de una red hidráulica son bien conocidos por los expertos en la materia de la ingeniería de sistemas de agua. Por lo general consisten en, a partir del estado del tanque en una referencia de tiempo y los valores de las variables de control, usar la ley física de la hidráulica para calcular los valores de todas las variables de estado en la misma referencia de tiempo, y luego calcular los valores del estado del tanque en la siguiente referencia temporal. Entonces es posible calcular los valores predichos de las variables de estado en todas las referencias temporales a partir de los valores iniciales de las variables de estado y los valores de las variables de control. Los procedimientos para calcular los valores predichos de las variables de estado en una referencia temporal aparecen desvelados principalmente en O. Piller, "Modeling the behavior of a network-Hydraulic analysis and sampling procedures for parameter estimation". PhD thesis in Applied Mathematics from the Mathematics and Computer Science Doctoral School at the University of Bordeaux (PRES), 1995, 288 páginas, Talence, Francia.

El procedimiento comprende además una etapa 330 de cálculo de valores residuales del subconjunto de las variables de estado como una diferencia entre predicciones y observaciones en las segundas referencias temporales. En caso de que las segundas referencias temporales y las primeras referencias temporales no sean idénticas, los valores de las observaciones pueden calcularse en las segundas referencias temporales, por ejemplo, construyendo una función continua de cada observación en la primera referencia temporal y seleccionando valores en las segundas referencias temporales deseadas. Dicha función continua se puede construir de varias maneras. Por ejemplo, puede construirse utilizando una función de interpolación o una función de regularización. Aquí se adaptan varias funciones, por ejemplo, una función de convolución del núcleo gaussiano. Esta etapa consiste en calcular las diferencias entre los valores predichos y observados de las variables de estado, para variables de estado para las cuales hay observaciones disponibles. Las observaciones de las variables de estado pueden ser observaciones en bruto u observaciones preprocesadas. Los valores observados de las variables de estado pueden, por ejemplo, preprocesarse mediante suavizado, falta de imputación de datos, etc. para limpiar la señal en los casos en que el sistema está correctamente modelado y calibrado, y las entradas y salidas de la red están correctamente definidas, entonces, si no hay una anomalía real, las observaciones y predicciones de las variables de estado son muy cercanas y los valores residuales están muy cerca de los valores residuales en la calibración del modelo. Por el contrario, los valores residuales pueden ser importantes, es decir, se produce una anomalía, como por ejemplo cuando un consumo espacial de agua es incorrecto. La función objetivo se usa para calcular la importancia de los valores residuales. La importancia del valor residual puede calcularse, por ejemplo, como una norma ponderada o un promedio ponderado del valor absoluto de cada valor residual. Los pesos asociados con las diferentes observaciones/residuos se pueden definir de acuerdo con la naturaleza de la observación (flujo, presión, etc.), la marca y la precisión del sensor, la hora del día, la varianza empírica de la medición. Más generalmente, el peso puede ser definido automáticamente, o definido por un usuario, según una indicación de fiabilidad de la medición.

El procedimiento 250 comprende además una etapa 340 de verificar si los valores residuales satisfacen un criterio de interrupción. Esta etapa consiste en verificar si los valores residuales de las variables de estado son lo suficientemente pequeños como para considerar que los valores de las variables de control describen con precisión el comportamiento de la entidad. El criterio de interrupción puede ser, por ejemplo, una combinación de un umbral en el número de iteraciones, un umbral sobre la variación de los valores residuales, un umbral en el gradiente de la función objetivo con respecto al conjunto de parámetros.

Cuando el criterio de interrupción 340 no se cumple, el procedimiento 250 comprende una etapa de bucle 350 para cambiar los valores de los parámetros del problema (por ejemplo, el perfil de consumo) usando un procedimiento de descenso. En una realización, el procedimiento puede ser un algoritmo de gradiente, un algoritmo de gradiente conjugado, un algoritmo de Newton, un algoritmo de Levenberg-Marquardt u otro algoritmo del mismo tipo. Esta etapa 350 alimenta una nueva iteración de la etapa 320, utilizando el modelo hidráulico para calcular valores predichos con valores modificados de parámetros del problema.

Cuando se cumple el criterio de interrupción 340, el procedimiento 250 comprende una etapa 360 que es una prueba basada en un criterio de refinamiento. Mientras el valor de la función objetivo disminuye, otra etapa de refinamiento 370 se realiza en el bucle en la etapa 320.

El procedimiento 250 comprende además una etapa 370 para seleccionar un subconjunto de al menos un elemento (nodo o arco). La etapa examina el gradiente calculado como el producto escalar de un vector de sensibilidad de elemento (nodo o arco) por un vector de contribución de elemento. El procedimiento descarta los elementos de acuerdo con un criterio de selección, en función de la evaluación de sensibilidad. En una realización, los que tienen una sensibilidad positiva se descartan. Esta etapa 370 alimenta una nueva iteración de la etapa 320, utilizando el modelo hidráulico para calcular los valores predichos en los elementos seleccionados.

El procedimiento 250 comprende además una etapa 380 para proporcionar las salidas de los parámetros del modelo. Las salidas del modelo incluyen los parámetros ajustados, los valores sucesivos de los conjuntos de elementos y la función objetivo relacionada, y las características que se procesan en la etapa 260 tal y como se muestra en la figura 2. Las salidas del modelo incluyen, por lo tanto, información sobre diversas entidades, desde un nodo o arco específico

hasta un subconjunto de la red y toda la red.

La figura 4 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención.

5 El procedimiento 400 de la invención comprende etapas similares al procedimiento 200 de la técnica anterior, principalmente:

- la etapa 220 de parametrizar un modelo hidráulico del sistema de distribución de agua con un conjunto de valores de variables de control que caracterizan la red y su salida en los nodos;
- la etapa 230 de usar sensores en la red para adquirir observaciones de un subconjunto de variables de estado en las primeras referencias temporales;
- 10 - la etapa 250 de cambiar el conjunto de valores de las variables de control usando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado;
- la etapa 270 de realizar una clasificación de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control.

15 En una serie de realizaciones de la invención, el procedimiento también puede comprender una o más de las otras etapas del procedimiento 200:

- la etapa 280 de detección de eventos;
- la etapa 290 de mostrar eventos y sus características;
- la etapa 291 de enriquecer la base de datos de clases normales históricas;
- la etapa 292 de enriquecer la base de datos de clases históricas anormales;

20 El procedimiento 400 comprende además una etapa 410 de identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones.

25 El objetivo de la etapa 410 es reducir el número de entidades en las que los valores de las variables de control se cambian en la etapa 250, asegurando que, si ocurre una anomalía en el sistema de distribución de agua, se encuentra en las entidades seleccionadas en la etapa 410. De hecho, la carga computacional y la complejidad de la etapa 250 pueden aumentar enormemente en redes grandes con un alto número de entidades.

De este modo, al reducir el número de entidades para las cuales se modifican las variables de control, el procedimiento de la invención reduce la carga computacional de la detección de anomalías y produce sus resultados más rápido y al menos de manera igual de fiable que los procedimientos de la técnica anterior.

30 En una serie de realizaciones de la invención, la etapa 410 identifica además otros elementos, por ejemplo las variables de control que se han de cambiar, reduciendo de este modo adicionalmente el coste de cálculo de la etapa 250.

En una serie de realizaciones de la invención, la etapa 250 se realiza para una pluralidad de tipos de eventos, y la etapa 410 identifica posibles tipos de eventos para probar.

35 De acuerdo con diversas realizaciones de la invención, la identificación de las entidades en la que cambiar los valores de control, de los valores de control que se han de cambiar y de los eventos que se han de probar se puede realizar de muchas maneras diferentes. Por ejemplo, puede basarse en un análisis estadístico de los valores residuales de las variables de estado y una primera clasificación de entidades. También puede basarse en una comparación directa de valores de variables de estado, por ejemplo la velocidad del agua, con un umbral.

40 Más generalmente, en un procedimiento de acuerdo con la invención, se detecta una anomalía, si al menos una entidad se clasifica en un estado representativo de una anomalía en la etapa 270, o si se detecta un evento en la etapa 280. En una serie de realizaciones de la invención, algunos eventos son representativos de una anomalía, mientras que otros no lo son. De este modo, la etapa 290 de mostrar eventos y sus características comprende mostrar una anomalía a un operador, si al menos una entidad ha sido clasificada en un estado anormal. En otras realizaciones de la invención, también es posible visualizar o registrar, además de eventos, las clasificaciones en un estado normal. De este modo, un experto puede verificar esta visualización o registro de estados normales para garantizar que la clasificación se realice correctamente y que ninguna clasificación errónea lleve a perder una anomalía.

45 Se presentarán ejemplos de realizaciones de la invención, a modo de ejemplos no limitativos, en el resto de la descripción.

50 La figura 5 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, en el que la ejecución del ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico combinado con un aprendizaje automático para identificar la anomalía depende de la salida de una clasificación de anomalías utilizando un cálculo de los valores residuales de las variables de estado.

El procedimiento 500 comprende las etapas 220, 230, 250, 260 y 270 del procedimiento 400. En una serie de realizaciones de la invención, también puede comprender al menos una de las etapas 280, 290, 291, 292 del

procedimiento 400.

5 En este ejemplo de realizaciones, la etapa 410 de identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control se realiza utilizando un análisis estadístico de los valores residuales de las variables de estado. El uso de esta combinación permite configurar la parte iterativa directamente en una clase de anomalía
 10 previamente identificada con una entidad previamente identificada, y centrarse en anomalías importantes para que el procedimiento de detección sea más preciso y los operadores puedan usarlo más fácilmente. Las principales anomalías son las que producen discrepancias estadísticas entre los resultados del modelo y los datos medidos, que se detectan a través del análisis de residuos. El procedimiento iterativo confirma la presencia de una anomalía y brinda información más detallada. Luego se puede iniciar con menos frecuencia y con un menor número de instancias paralelas, ahorrando tiempo de cálculo.

La etapa 410 de identificación de al menos un objetivo comprende de este modo la etapa 320 de cálculo de los valores predichos, y la etapa 330 de cálculo de residuos, que se conocen a partir de la técnica anterior. Los valores residuales se pueden calcular para todas o una parte de las variables de estado para las que está disponible una observación.

15 La etapa 410 comprende además una etapa 510 de realizar una pluralidad de análisis estadísticos de valores residuales para al menos una entidad para una selección de referencias temporales. Esta etapa consiste en analizar la distribución de valores residuales para una entidad en cada referencia temporal en una ventana de tiempo. De acuerdo con diversas realizaciones de la invención, la ventana de tiempo puede cubrir la totalidad o un subconjunto de referencias temporales.

20 A través de la pluralidad de análisis estadísticos, se puede detectar una anomalía si los valores residuales no son consistentes con las observaciones históricas de los valores residuales. El uso de la distribución histórica de valores residuales permite detectar los comportamientos habituales e inusuales, incluyendo variabilidad y varias propiedades de los valores residuales, para su comparación con los valores residuales actuales dentro de una ventana de tiempo dada. Por ejemplo, la etapa 510 puede comprender calcular inferencias estadísticas (momento de orden n, promedio, desviación estándar, desviación media o mediana absoluta, etc.) utilizando una base de datos de valores residuales
 25 históricos para su comparación con las inferencias estadísticas correspondientes de los valores residuales actuales. Esto le da más robustez al procedimiento ya que toma en cuenta los valores pasados y los comportamientos observados.

30 En una serie de realizaciones de la invención, los valores residuales se transforman usando una transformada de Fourier antes del análisis estadístico. En una realización de la invención, realizar un análisis estadístico para una entidad comprende detectar si un valor residual excede un umbral predefinido para un número predefinido de etapas sucesivas. El umbral y el número de etapas de tiempo sucesivas se pueden predefinir en función de detecciones anteriores, para maximizar la relación de detección de anomalías cuando existe una anomalía (verdaderos positivos) y minimizar la relación de detección de una anomalía cuando no existe (falsas alarmas). El análisis puede usar pruebas estadísticas sobre las propiedades de los valores residuales y el cálculo de los valores P resultantes para calificar el
 35 resultado de las pruebas. Los umbrales de los valores P se pueden establecer para obtener un nivel de riesgo de obtener falsas alarmas. Establecer un valor dado implica esperar un porcentaje dado de falsas alarmas, es decir, rechazo incorrecto de suposiciones. En una realización, el análisis estadístico se realiza típicamente en valores residuales de presión o velocidad. Por ejemplo, un análisis del caudal para detectar una fuga puede realizarse en ventanas de unas pocas horas.

40 La etapa 410 comprende además una etapa 520 de clasificar al menos una entidad de la red en función de las reglas que se aplicarán a la salida de la etapa de análisis estadístico. Siguiendo la aplicación de las reglas, se da una clase a las entidades (por ejemplo, "fuga", "caída de presión", "problema de calidad del agua"). En una realización de la invención, la regla se basa en una comparación de los valores P dados por las pruebas estadísticas realizadas en la etapa 510, con un nivel de riesgo predefinido, dependiendo del nivel objetivo de sensibilidad del procedimiento. En una realización de la invención, la regla se basa en un algoritmo de clasificación previamente capacitado, permitiendo,
 45 en función de la simulación pasada, analizar el vector actual de valores residuales y dar la clase actual de la entidad. En una realización de la invención, hay dos clases disponibles, una representativa de estado normal y la otra representativa de uno anormal. En otras realizaciones de la invención, las clases dan una naturaleza más precisa de la anomalía, como "caída anormal de presión", "problema con la calidad del agua", etc.

50 El procedimiento 500 comprende además una etapa 530 para determinar la siguiente etapa de acuerdo con las clases actuales. Si una de las clases corresponde a un estado anormal (uno de los posibles estados anormales) o si el estado se detecta como periférico, el procedimiento pasa a 250. Si todas las clases se califican como normales, el procedimiento pasa a 540.

55 Si una de las clases corresponde a una clase anormal, la etapa 250 puede adaptarse a las entidades que se han clasificado como anormales. En una serie de realizaciones de la invención, solo las variables de control de las entidades clasificadas como anormales pueden modificarse inicialmente en la etapa 250. En otras realizaciones de la invención, controlar las variables de las entidades que se clasifican como anormales, así como las variables de control de entidades adyacentes pueden modificarse inicialmente en la etapa 250. También es posible asociar una clase en la salida de la etapa 520 para controlar variables para modificar o eventos para probar en la etapa 250. Por lo tanto,

5 la etapa 250 se lanza para un número reducido de posibles anomalías, correspondiente a las anomalías esperadas, en subconjuntos específicos de nodos y arcos. Este procedimiento iterativo mejor sintonizado y enfocado geográficamente da como resultado una caracterización de anomalía más rápida y precisa (ubicación, intensidad). El mejor ajuste se debe, en particular, a las condiciones de contorno más fuertes establecidas para los problemas, mientras que en la etapa 220, el modelo hidráulico es menos limitado.

10 En la salida de la etapa 270, el procedimiento 500 comprende una etapa 560 de enriquecer una base de datos de residuos históricos con estados normales/anormales. Esta etapa consiste en asociar los residuos con una etiqueta "normal" si no se ha identificado ninguna anomalía en la etapa 270, una etiqueta "anormal" o una etiqueta representativa de un evento, si se ha detectado un evento en la salida de la etapa 270. Esto permite enriquecer la base de datos que asocia valores residuales con un estado normal, un estado anormal o un evento, y permite obtener una clasificación adicional de los valores residuales en las ejecuciones posteriores del procedimiento, mejorando de este modo su eficiencia en el tiempo.

La etapa 540 consiste en enriquecer una base de datos de residuos con estados normales. De manera similar a la etapa 560, esto permite mejorar la clasificación de las entidades en la etapa 520 a lo largo del tiempo.

15 El procedimiento comprende además una etapa 550 que enriquece la base de datos histórica de clase normal, en caso de que no se cumpla el criterio 530. Esta etapa consiste en ejecutar diferentes ajustes iterativos paso a paso de propiedades y variables de control de un modelo hidráulico, y extraer vectores de características correspondientes a los diferentes tipos de anomalías. Estos vectores se consideran directamente como representantes de clase normales, posiblemente con diferentes contextos, y almacenados en la base de datos de clase normal para su uso posterior para capacitar el algoritmo de aprendizaje automático.

EJEMPLO DE REALIZACIÓN: DETECCIÓN DE UNA DISMINUCIÓN ANORMAL DE LA CONCENTRACIÓN DE CLORO RESIDUAL

A modo de ejemplo no limitativo del procedimiento 500, a continuación se describe un escenario de detección de una disminución anormal de la concentración de cloro residual.

25 Un sistema de distribución de agua está equipado con sensores que miden parámetros representativos de la calidad del agua, además de sensores hidráulicos. En particular, comprende sensores de la concentración de cloro residual (denominada RCC a continuación).

30 Un modelo hidráulico de la red ha sido calibrado durante una fase de capacitación, así como un modelo de la evolución de la RCC. En este ejemplo, los valores predichos, observados y residuales son valores de la RCC. Durante una fase de capacitación, se aplica un módulo de aprendizaje a los valores residuales de la RCC, para determinar los parámetros de los valores residuales representativos de una RCC normal.

35 Una vez que se calculan los valores residuales de RCC, el análisis estadístico compara los valores que son característicos de la distribución de los valores residuales con los valores históricos utilizados durante la fase de capacitación. Si los valores residuales que tienen al menos una propiedad de distribución tienen un valor anormal, un valor absoluto por encima de un umbral, para un gran número de tiempos sucesivos de referencia, un subconjunto de la red se clasifica previamente en una clase de "disminución anormal de RCC" con un índice de fiabilidad.

40 El ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico se inicia con el subconjunto específico de la red y la configuración correspondiente a la anomalía de la calidad del agua, en particular la evolución de la RCC. Las propiedades y las variables de control incluyen aquí parámetros específicos del modelo de cloro, como las constantes cinéticas para agua a granel en las diversas fuentes y las constantes cinéticas para cada material de tubería.

45 Si no se encontró previamente ningún caso de distribución anormal de RCC en la red, el procesamiento de los resultados a la salida de la resolución del modelo inverso se puede realizar utilizando una detección de anomalías o una agrupación aplicada a las constantes cinéticas. Al usar la detección de anomalías, el estado se compara con los estados históricos. Se declara como anomalía si la distancia está por encima de un umbral. Al menos un valor se clasifica en una clase de "disminución anormal de RCC" con un índice de fiabilidad.

50 La información sobre la anomalía y su localización se puede presentar a un operador a través de una interfaz gráfica. Esta realización es ventajosa, ya que proporciona robustez adicional a la detección de anomalías. De hecho, se detectará una anomalía si se genera primero mediante el análisis de residuos y luego se confirma y localiza gracias al modelo inverso, que se ha sintonizado más específicamente.

55 La figura 6 muestra un ejemplo de procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, en la que la ejecución de un ajuste iterativo paso a paso de propiedades y variables de control de un modelo hidráulico combinado con un aprendizaje automático para identificar la anomalía depende de la salida de una clasificación de anomalías usando una combinación de reglas aplicadas a las variables de estado.

El procedimiento 600 comprende las etapas 220, 230, 250, 260 y 270 del procedimiento 400. En una serie de realizaciones de la invención, también puede comprender al menos una de las etapas 280, 290, 291, 292 del procedimiento 400.

5 En una serie de realizaciones de la invención, la etapa 410 del procedimiento 600 comprende una etapa 610 de realizar una serie de pruebas sobre observaciones de acuerdo con una combinación de reglas, sin usar un modelo hidráulico, y la etapa 620 de clasificar un nodo, en función de la salida de la etapa 610.

10 En la etapa 610, las observaciones de las variables de estado se prueban para determinar si hay anomalías y en qué punto, antes de parametrizar el procedimiento iterativo. Se realizan varias pruebas y sus resultados se combinan utilizando reglas para clasificar entidades en los diferentes estados predefinidos que se han predefinido. En una
 15 realización, el procedimiento compara todos los valores medidos de las variables con umbrales predefinidos que se han definido en función de la experiencia de los operadores. Si el valor de una variable excede el umbral, las anomalías correspondientes se seleccionan en un área que está predefinida en función de las ubicaciones de los diferentes sensores. En otra realización, se detecta una anomalía si el número de etapas de tiempo de observación excede un umbral dado. En una realización, el número de etapas de tiempo que exceden el umbral para elevar la anomalía se da a través del cálculo de probabilidad binomial y un nivel de riesgo dado. En otra realización, las nuevas variables se construyen en función de observaciones, para describir aspectos particulares que podrían estar sujetos a anomalías. En una realización, la pendiente de la regresión lineal en el tiempo se puede calcular y comparar con los umbrales para detectar una evolución anormal. En otra realización, se puede calcular un determinado percentil en observaciones (por ejemplo, 5 % o 50 %) y compararlo con umbrales predefinidos. En otra realización, los umbrales se pueden
 20 calcular en función de la combinación lineal de valores medios y desviaciones estándar observadas pasadas en las observaciones para tener en cuenta la variabilidad habitual de las señales. En una realización, los umbrales diferentes pueden basarse típicamente en 2 y 3 desviaciones estándar alrededor de los valores medios.

25 En una realización, cada resultado de la prueba se puede puntuar y las puntuaciones se pueden combinar con una combinación multilineal. Las puntuaciones pueden ser solo 0 y 1, pero también pueden ser más precisos, para reflejar diferentes niveles de sospecha, por ejemplo, vinculados con diferentes niveles de riesgo (con valores P) o diferentes umbrales. La combinación de puntuaciones por lo tanto proporciona un índice indicativo de la posibilidad de que una entidad esté en un estado anormal, aumentando así la robustez. En una realización, la combinación puede hacer uso de pesos, que se pueden ajustar en las fases de capacitación, para dar más importancia a una prueba en particular. En una realización, una de las pruebas puede ser más fiable, pero otras pueden reunirse para ser más fiables en
 30 conjunto. Esta combinación también se aplica a los resultados de las pruebas entre varias entidades. En una realización, la combinación se usa para evaluar si el comportamiento anormal es similar en muchas entidades, por lo tanto, son los resultados de un controlador común y no una anomalía real. Esta combinación tiene como objetivo aumentar la solidez de la inicialización del procedimiento iterativo. Usando este procedimiento, se puede probar una amplia variedad de reglas, en una serie de diferentes variables. Tal y como se explica a continuación, se pueden identificar las variables y reglas relevantes, adaptadas y mejoradas para una red dada en las etapas 540, 550 y 560. En una realización, las pruebas que se han de aplicar, las reglas y la combinación de puntuaciones se aplican de
 35 manera diferente en grupos de áreas. Estos grupos de áreas pueden ser definidos por el usuario o por un procedimiento de agrupación en función de datos históricos.

40 La etapa 410 del procedimiento 600 comprende además una etapa 620 de clasificación de una entidad que es muy similar a la etapa 520 del procedimiento 500, salvo por que se aplica a la salida de la combinación de las reglas 610, en lugar de la producción de estadísticas sobre valores residuales 510. De manera similar a la etapa 520, la etapa 620 se puede capacitar usando un algoritmo de aprendizaje automático, o adaptarse específicamente para clasificar las entidades de acuerdo con el resultado de la etapa 610. De forma similar, el procedimiento 600 comprende una etapa 630 para detectar o no una anomalía, que es muy similar a la etapa 530 del procedimiento 500, salvo por que se aplica a la salida de la etapa 620.

50 Por lo tanto, la etapa 250 se inicia con un número reducido de configuraciones, correspondiente a las anomalías esperadas, en subconjuntos específicos de nodos y arcos. Este procedimiento iterativo mejor sintonizado y enfocado geográficamente da como resultado una caracterización de anomalía más rápida y precisa (ubicación, intensidad). El mejor ajuste se debe, en particular, a las condiciones de contorno más fuertes establecidas para los problemas, mientras que en la etapa 220, el modelo hidráulico es menos limitado.

55 En una serie de realizaciones de la invención, el procedimiento 600 comprende además la etapa 640 de enriquecer la base de datos histórica de reglas con estados normales, y la etapa 660 de enriquecer la base de datos histórica de reglas con estados normales y anormales. Estas etapas son muy similares a las etapas 540 y 560 respectivamente, en cambio, se enriquece una base de datos de asociación de reglas y estados normales/anormales, en lugar de una base de datos de reglas de asociación de valores residuales con estados normales/anormales. Este enriquecimiento permite, durante varias ejecuciones del procedimiento, capacitar las reglas de detección previa de anomalías a una red de distribución de agua dada. De este modo, incluso sin un conocimiento a priori de las reglas sobre las variables de estado que son representativas o no de un estado anormal, el procedimiento puede adaptarse automáticamente a la red y adoptar las reglas más relevantes. En una serie de realizaciones de la invención, el sistema se inicializa con reglas arbitrarias: por ejemplo, el 1 % de los nodos que tienen el flujo de agua más importante se clasificarán como
 60 anormales. La ejecución de la etapa 250 permite determinar aquellos que realmente están en un estado anormal, y la

etapa 660 puede enriquecer la base de datos con el conocimiento del nodo que estaba en un estado normal real o anormal. Durante varias ejecuciones del procedimiento, las reglas y la clasificación en las etapas 610 y 620 adquieren un conocimiento suficiente de eventos pasados para cambiar de reglas arbitrarias a reglas de detección previa de eventos. En una serie de realizaciones de la invención, el procedimiento 600 también comprende la etapa 550 del procedimiento 500.

EJEMPLO DE REALIZACIÓN: DETECCIÓN DE UN INCREMENTO ANORMAL DE LA DEMANDA DE AGUA UTILIZANDO UNA DETECCIÓN PREVIA EN FUNCIÓN DE UNA COMBINACIÓN DE REGLAS

A modo de ejemplo no limitativo del procedimiento 600, a continuación se describe ahora un escenario de detección de un aumento anormal de la demanda de agua.

Este ejemplo demuestra la capacidad de un procedimiento 600 de acuerdo con la invención para detectar un aumento anormal de la demanda de agua, posiblemente vinculado a una fuga, con una detección previa en función de una combinación de reglas.

Este ejemplo describe la detección de un aumento anormal de la demanda de agua, que puede estar vinculado a una fuga. Las series temporales observadas de caudales se combinan para formar demandas de agua para áreas determinadas. Se calculan diferentes percentiles en estas nuevas series temporales y se comparan con umbrales en función de combinaciones lineales de desviaciones estándar observadas pasadas. En función de las ejecuciones anteriores del procedimiento, se ha aprendido un número límite de etapas de tiempo sucesivas para los cuales los percentiles deben estar por encima del umbral para detectar previamente una anomalía. Al utilizar el aprendizaje automático, este número límite de etapas de tiempo sucesivas se calcula como un número óptimo para evitar posibles anomalías y no tener demasiadas detecciones previas. Supongamos que para 3 áreas dadas, los percentiles están por encima de los umbrales en más etapas de tiempo que el número límite de etapas de tiempo que se aprendió para cada área dada. Para todas las demás áreas, los percentiles están por debajo de los umbrales en el número requerido de etapas de tiempo. Para ellos, la puntuación dada es 0. Para cada una de las 3 áreas que activan los umbrales, se da una puntuación de 1 y luego se aplica la regla de combinación: Para cada área, si su propia puntuación es 1, el porcentaje de otras áreas con una puntuación igual a 0 se compara con un umbral predefinido. En el caso actual, este porcentaje está por encima del umbral para las 3 áreas, ya que muchas otras áreas tienen una puntuación de 0. Por lo tanto, las 3 áreas se detectan previamente como anormales. La detección previa de anomalías en 3 áreas en lugar de una evita una falsa alarma en función de posibles cambios de demanda de agua que se producen al mismo tiempo en muchas áreas. Estos posibles cambios pueden deberse a la estacionalidad o un evento común específico que no debe ser señalado a la atención de los operadores. Esto aumenta la robustez del procedimiento al disminuir el índice de falsas alarmas.

El ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico se inicia con el subconjunto específico de estas 3 áreas de la red y la configuración correspondiente al aumento de la demanda de agua. Las propiedades y las variables de control están configuradas para detectar fugas en las 3 áreas dadas, como subconjuntos de nodos en los que se asigna una demanda anormal (su número disminuye a través del procedimiento de iteración) y sus perfiles de tiempo de consumo de agua.

La información sobre las anomalías y sus localizaciones se puede presentar a un operador a través de una interfaz gráfica. Esta realización es ventajosa, ya que proporciona robustez adicional a la detección de anomalías. De hecho, se detectará una anomalía si primero se plantea una combinación de reglas sobre observaciones y se confirma mediante el modelo inverso, más específicamente sintonizado.

La figura 7 muestra un ejemplo de un procedimiento para detectar anomalías en un sistema de distribución de agua en una serie de realizaciones de la invención, utilizando el aprendizaje estadístico para detectar y caracterizar previamente una anomalía y el uso de un ajuste iterativo paso a paso de propiedades y variables de control de un modelo hidráulico, agregado al aprendizaje automático para identificar la anomalía. La forma de usar esta combinación permite configurar la parte iterativa directamente en una clase de anomalía previamente identificada, con una entidad previamente identificada y centrándose en las principales anomalías para hacer que el procedimiento de detección sea más preciso y fácil de usar por los operadores. El procedimiento iterativo confirma la presencia de una anomalía y brinda información más detallada.

El procedimiento 700 comprende las etapas 220, 230, 250, 260 y 270 del procedimiento 400. En una serie de realizaciones de la invención, también puede comprender al menos una de las etapas 280, 290, 291, 292 del procedimiento 400.

En una serie de realizaciones de la invención, la etapa 410 del procedimiento 600 comprende una etapa 710 de realizar una serie de pruebas sobre observaciones de acuerdo con una combinación de reglas, sin usar un modelo hidráulico, y una etapa 720 de clasificar un nodo, en función de la salida de la etapa 710.

En la etapa 710, las observaciones de las variables de estado se prueban para determinar si hay anomalías y en qué punto, antes de parametrizar el procedimiento iterativo. El procedimiento simula los valores esperados de las observaciones, en función de sus valores pasados y controladores adicionales. En una serie de realizaciones de la invención, se capacitan varios modelos sobre datos pasados, y se selecciona el modelo que produce la mejor

predicción de datos pasados. En una realización, el procedimiento contiene modelos basados en análisis de series temporales como la técnica de promedio móvil autorregresivo (ARMA) o la técnica de promedio móvil integrado autorregresivo estacional (SARIMA). Los modelos ARMA y SARIMA aparecen desvelados principalmente en Box, George; Jenkins, Gwilym (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day. En una

5 realización, el procedimiento contiene modelos basados en algoritmos de aprendizaje automático, como bosques aleatorios o redes neuronales artificiales. En una realización, el procedimiento incluye modelos que usan valores pasados o actuales de otras variables para una observación dada de la variable. Los controladores adicionales son variables o parámetros que influyen en las observaciones. En una realización, se incluyen los volúmenes de temperatura y precipitación. En otra realización, se incluye el calendario y los períodos de tiempo con eventos

10 específicos. Los valores simulados para las observaciones se comparan con las observaciones reales a través del cálculo de valores residuales. Los valores residuales se prueban con diferentes pruebas y sus resultados se combinan mediante reglas para clasificar entidades en los diferentes estados predefinidos. En una realización, el procedimiento prueba todos los valores residuales en comparación con umbrales predefinidos en función de la experiencia de los operadores. Si los valores residuales exceden los umbrales, las anomalías correspondientes se seleccionan en un

15 área que se define en función de la ubicación de los diferentes sensores. En otra realización, las anomalías se marcan si demasiadas etapas de tiempo de valores residuales exceden un umbral dado, para una ventana de tiempo dada. Esta ventana de tiempo depende de la naturaleza de las variables y las anomalías que se han de detectar. En una realización, el período de tiempo puede ser de 6 horas. En una realización, el número de etapas de tiempo que exceden el umbral para elevar la anomalía se da a través del cálculo de probabilidad binomial y un nivel de riesgo

20 dado. En otra realización, las nuevas variables se construyen en función de valores residuales, para describir aspectos particulares que podrían estar sujetos a anomalías. En una realización, la suma acumulada de valores residuales a lo largo del tiempo puede calcularse y compararse con los umbrales para detectar una evolución anormal. En otra realización, la distribución de los valores residuales se prueba con un procedimiento estadístico, para su comparación con la distribución pasada de residuos.

25 Si una prueba revela un cambio significativo en la distribución, se puede marcar una anomalía. En otra realización, los umbrales se pueden calcular en función de las desviaciones estándar observadas pasadas en las observaciones para tener en cuenta la variabilidad habitual de las señales. En una realización, los números de desviaciones estándar que se utilizarán pueden ser típicamente 2 y 3. En una realización, cada resultado de la prueba se puede puntuar y las puntuaciones se pueden combinar con una combinación multilineal. Las puntuaciones pueden ser solo 0 y 1, pero

30 también pueden ser más precisos, para reflejar diferentes niveles de sospecha, por ejemplo, vinculados con diferentes niveles de riesgo (con valores P) o diferentes umbrales. La combinación de puntuaciones por lo tanto da el nivel general de sospecha en cada entidad, aumentando la robustez. En una realización, la combinación puede hacer uso de pesos, que se pueden ajustar en las fases de capacitación, para dar más importancia a una prueba en particular. En unas realizaciones, una de las pruebas puede ser más fiable, pero otras pueden reunirse para ser más fiables en

35 conjunto. Esta combinación también se aplica a los resultados de las pruebas entre varias entidades. En una realización, la combinación se usa para evaluar si el comportamiento anormal es similar en muchas entidades, por lo tanto, son los resultados de un controlador común y no una anomalía real. Esta combinación tiene como objetivo aumentar la solidez de la inicialización del procedimiento iterativo. Los algoritmos de aprendizaje automático permiten aprender qué variables de estado, en qué entidad, son representativas de una anomalía en una red determinada. Más

40 generalmente, la fiabilidad de los resultados de las pruebas se puede mejorar utilizando procedimientos de aprendizaje que asocian umbrales en las observaciones a la presencia o ausencia de una anomalía; el principio de adaptación del umbral es similar a la adaptación de las reglas, salvo por que se aplica directamente en umbrales de observaciones, en lugar de aplicarse a las reglas.

45 La etapa 410 del procedimiento 700 comprende además una etapa 720 de clasificación de una entidad que es muy similar a la etapa 520 del procedimiento 500, salvo por que se aplica a la salida de las pruebas de la etapa 710, en lugar de la producción de estadísticas sobre valores residuales 510. De manera similar a la etapa 520, la etapa 720 se puede capacitar con un algoritmo de aprendizaje automático, o adaptarse específicamente para clasificar las entidades de acuerdo con el resultado de la etapa 710. De forma similar, el procedimiento 600 comprende una etapa

50 730 para detectar o no una anomalía, que es muy similar a la etapa 530 del procedimiento 500, salvo por que se aplica a la salida de la etapa 720.

Por lo tanto, la etapa 250 se inicia con un número reducido de anomalías detectadas previamente, correspondiente a las anomalías esperadas, en subconjuntos específicos de nodos y arcos. Este procedimiento iterativo mejor sintonizado y enfocado geográficamente da como resultado una caracterización de anomalía más rápida y precisa (ubicación, intensidad). El mejor ajuste se debe, en particular, a las condiciones de contorno más fuertes establecidas

55 para los problemas, mientras que en la etapa 220, el modelo hidráulico es menos limitado.

En una serie de realizaciones de la invención, el procedimiento 700 comprende además la etapa 740 de enriquecer la base de datos histórica de observaciones con estados normales, y la etapa 760 de enriquecer la base de datos histórica de observaciones con estados normales y anormales. Estas etapas son muy similares a las etapas 540 y 560 respectivamente, salvo por el hecho de que se enriquece una base de datos de asociación de observaciones y estados

60 normales/anormales, en lugar de una base de datos de observaciones de asociación de valores residuales con estados normales/anormales. En una serie de realizaciones de la invención, el procedimiento 600 también comprende la etapa 550 del procedimiento 500. De este modo, después de varias ejecuciones del procedimiento, los algoritmos de aprendizaje automático permiten identificar las variables de estado, entidades y umbrales que son representativos de

una anomalía.

EJEMPLO DE REALIZACIÓN: DETECCIÓN DE UN INCREMENTO ANORMAL DE LA DEMANDA DE AGUA UTILIZANDO UNA DETECCIÓN PREVIA EN FUNCIÓN DE UMBRALES DE OBSERVACIONES

5 Mediante un ejemplo no limitativo del procedimiento 700, a continuación se describe un escenario de detección de un aumento anormal de la demanda de agua.

De manera similar al ejemplo anterior, este ejemplo demuestra la capacidad de un procedimiento de acuerdo con la invención para detectar un aumento anormal de la demanda de agua, posiblemente vinculado a una fuga. Sin embargo, este ejemplo, basado en el procedimiento 700, utiliza una comparación directa de observaciones con umbrales en lugar de una combinación de reglas.

10 Este ejemplo describe la detección de un aumento anormal de la demanda de agua, que puede estar vinculado a una fuga. Los caudales observados se combinan para formar demandas de agua para áreas determinadas. Estas demandas de agua se simulan con el uso de demandas de agua pasadas y análisis de series de tiempo. Las variables simuladas se utilizan para calcular valores residuales con observaciones. Se calculan las sumas acumuladas de valores residuales. Se comparan con los umbrales. Estos umbrales se basan en las variaciones observadas de las sumas acumuladas de valores residuales en el pasado. Significa que usan la medición de las variabilidades, aquí en función de medias calculadas y desviaciones estándar. En este ejemplo, para 2 áreas en toda la red, la suma acumulada de valores residuales excede los umbrales dados. Esto significa que el comportamiento esperado en términos de demanda de agua no se observa en estas 2 áreas, como los valores residuales son lo suficientemente grandes, en un número suficiente de etapas de tiempo, para decir que hay una desviación estadística. Se genera una detección previa del aumento anormal de la demanda de agua.

20 El ajuste iterativo paso a paso de las propiedades y las variables de control de un modelo hidráulico se inicia con el subconjunto específico de la red y la configuración correspondiente al aumento de la demanda de agua. Las propiedades y las variables de control están configuradas para detectar fugas en las 2 áreas dadas, como subconjuntos de nodos en los que se asigna una demanda anormal (su número disminuye a través del procedimiento de iteración) y sus perfiles de tiempo de consumo de agua.

25 La información sobre las anomalías y sus localizaciones se puede presentar a un operador a través de una interfaz gráfica. Esta realización es ventajosa, ya que proporciona robustez adicional a la detección de anomalías. De hecho, se detectará una anomalía si primero se plantea una combinación de reglas sobre observaciones y se confirma mediante el modelo inverso, más específicamente sintonizado.

30 La Figura 8 muestra un ejemplo de una presentación de la localización de anomalías a un operador.

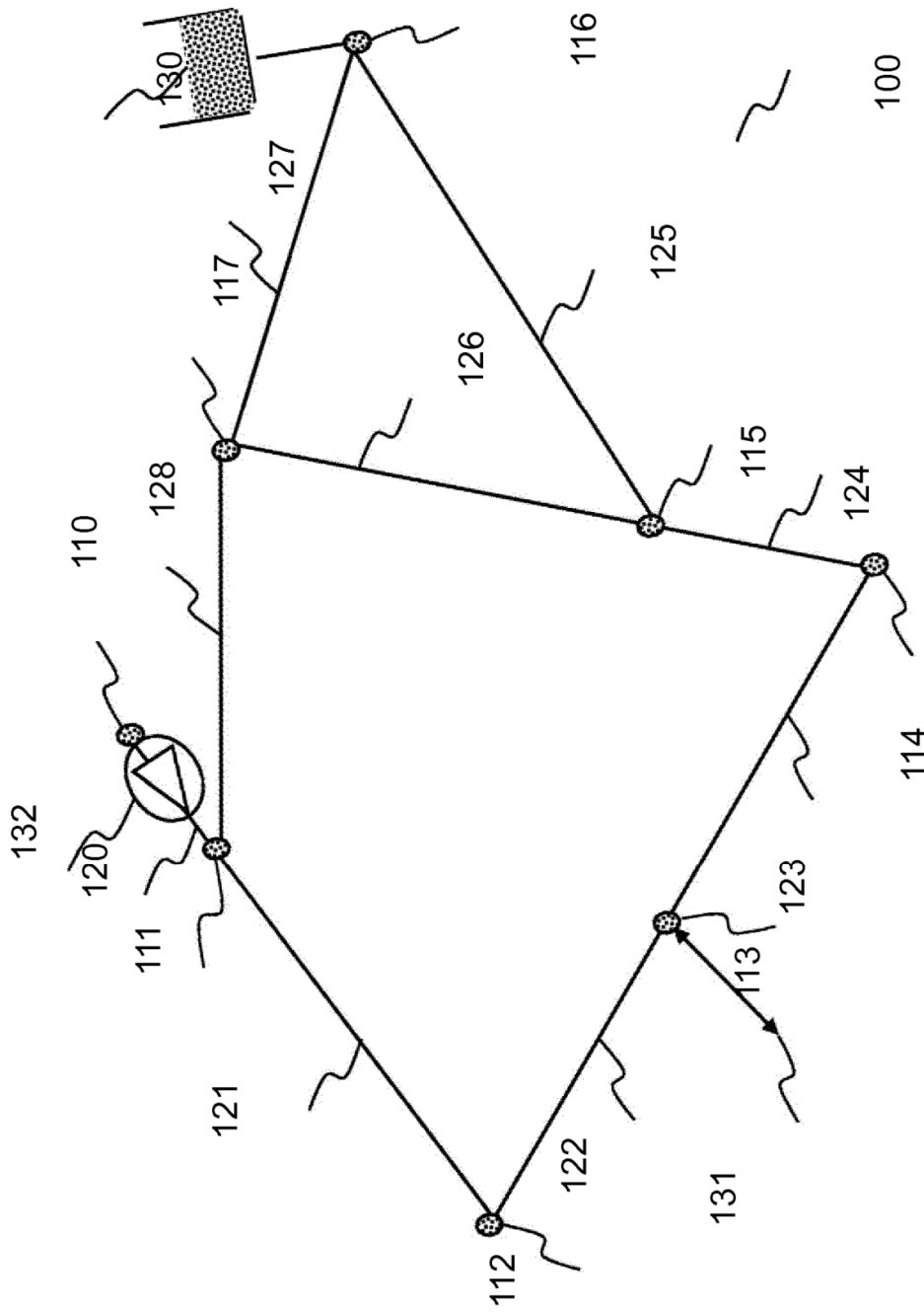
35 Un mapa 800 de la red de un sistema de distribución de agua se muestra a un operador a través de un dispositivo de visualización. El mapa contiene líneas que representan tuberías del sistema, el ancho de las líneas es representativo del diámetro de las tuberías. El sistema contiene tuberías grandes, como las tuberías 810 y 811, así como tuberías más pequeñas, tales como las tuberías 820 y 821. Tres fugas 830, 831, 832 se han localizado en la red y se representan mediante círculos grandes. Se puede insertar información sobre la importancia relativa de las fugas. Por ejemplo, el diámetro de los círculos puede aumentar con la importancia de la fuga. También es posible representar solo la fuga que se considera la más grave o mostrar información adicional sobre la fuga, por ejemplo, una lista de fugas pasadas que tenían las características más cercanas.

40 Los ejemplos descritos anteriormente se dan como ilustraciones de realizaciones de la invención. No limitan de ninguna manera el ámbito de la invención que se define mediante las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento (400) para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho procedimiento:
 - 5 - parametrizar (220) un modelo hidráulico (211) del sistema de distribución de agua con un conjunto de valores (210) de variables de control que caracterizan la red y su salida en los nodos;
 - utilizar sensores en la red para adquirir (230) observaciones de un subconjunto de variables de estado en primeras referencias temporales;
 - identificar (410) al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones;
 - 10 - cambiar (250) el conjunto de valores de las variables de control utilizando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado;
 - realizar una clasificación (270) de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control.
2. El procedimiento según la reivindicación 1, que comprende además identificar al menos una variable de control que se ha de modificar en función de al menos dichas observaciones.
3. El procedimiento según una de las reivindicaciones 1 y 2, en el que identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende:
 - 20 - utilizar el modelo hidráulico para calcular (320) valores predichos de un conjunto de variables de estado que caracterizan al menos una velocidad y presión del agua en los nodos, asociándose dichos valores predichos a segundas referencias temporales;
 - calcular (330) valores residuales del subconjunto de las variables de estado como una diferencia entre los valores predichos y las observaciones en las segundas referencias temporales;
 - realizar (510) un análisis estadístico de valores residuales en una entidad de la red en una selección de las referencias temporales;
 - 25 - clasificar (520) las entidades de la red en función de reglas aplicadas a la salida del análisis estadístico.
4. El procedimiento según la reivindicación 3, en el que dicha al menos una entidad objetivo se identifica clasificándose en un estado anormal en función de la comparación de valores residuales y un umbral predefinido.
5. El procedimiento según una de las reivindicaciones 1 y 2, en el que identificar al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones comprende identificar un arco de la red en el que el valor de una observación de la velocidad del agua excede un umbral en una entidad objetivo.
6. El procedimiento según una de las reivindicaciones 1 a 5, en el que el cambio (250) del conjunto de valores de las variables de control se realiza para al menos un tipo de evento, y las variables de control que se han de modificar se basan en dicho al menos un tipo de evento.
7. El procedimiento según la reivindicación 6, en el que se prueba una pluralidad de tipos de evento, y se realiza una instancia de cambio (250) del conjunto de valores de variables de control para cada tipo de evento.
8. El procedimiento según una de las reivindicaciones 6 y 7, cuando dependen de la reivindicación 3, en el que al menos cambiar el conjunto de valores de las variables de control se realiza para dicho al menos un tipo de evento elegido usando la clasificación (520) de las entidades de la red en función de las reglas aplicadas a la salida del análisis estadístico.
9. El procedimiento según una de las reivindicaciones 1 a 8, que comprende:
 - detectar (530) una anomalía en función de la salida de la identificación (410) de al menos una entidad objetivo en la que cambiar los valores de las variables de control en función de al menos dichas observaciones;
 - 45 - si no se detecta ninguna anomalía, enriquecer una base de datos de estados normales de las entidades (540);
 - si se detecta una anomalía:
 - oCambiar (250) el conjunto de valores de las variables de control utilizando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado;
 - oRealizar la segunda clasificación (270) de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control.
10. El procedimiento según una de las reivindicaciones 1 a 9, en el que cambiar el conjunto de valores de las variables de control utilizando un ajuste paso a paso de las variables de control y un criterio de interrupción en función de valores residuales de las variables de estado comprende:
 - A) cambiar el conjunto de valores de las variables de control (310);
 - B) utilizar el modelo hidráulico para calcular (320) valores predichos de un conjunto de variables de estado que

- caracterizan al menos la velocidad y la presión del agua en los nodos en las referencias temporales;
- 5 - C) calcular los valores residuales (330) del conjunto de variables de estado como una diferencia entre los valores predichos y las observaciones en las referencias temporales;
- D) si dichas diferencias cumplen un criterio de interrupción (340), pasar a la etapa F);
- E) si no, cambiar el conjunto de valores de las variables de control (350) y volver a la etapa B);
- F) si dichas diferencias no cumplen un criterio de refinamiento (360), seleccionar un subconjunto de la red (370) en que calcular los valores predichos, volviendo a la etapa B).
- 10 11. El procedimiento según la reivindicación 10, en el que el criterio de refinamiento comprende el cálculo de valores de una de al menos una función de cuadrado y una función objetivo bayesiana, y la selección y modificación de variables de control se determina mediante un algoritmo de Levenberg-Marquardt.
12. El procedimiento según una de las reivindicaciones 1 a 10, en el que la realización de una clasificación de al menos una entidad de la red en un estado de acuerdo con el conjunto de variables de control se realiza mediante un algoritmo de aprendizaje automático previamente capacitado.
- 15 13. El procedimiento según una de las reivindicaciones 1 a 12, en el que las variables de control comprenden variables escalares que caracterizan la topología y la topografía de la red, y variables basadas en el tiempo que caracterizan las entradas y salidas de la red que tienen al menos un valor en cada referencia temporal.
14. El procedimiento según una de las reivindicaciones 1 a 13, en el que las variables de estado caracterizan aún más la presión.
- 20 15. Un sistema para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho sistema:
- sensores de al menos velocidad y presión del agua en un subconjunto de nodos de la red;
 - un dispositivo informático que comprende un procesador;
 - enlaces de comunicación entre sensores y el dispositivo informático;
 - un medio de almacenamiento;
- 25 en el que el dispositivo informático está configurado para:
- recuperar un conjunto inicial de valores de variables de control que caracterizan la red y su salida en los nodos respecto del medio de almacenamiento y utilizarlo para parametrizar un modelo hidráulico del sistema de distribución de agua;
 - utilizar enlaces de comunicación entre sensores en la red para adquirir observaciones de un subconjunto de las
- 30 variables de estado, teniendo dichas observaciones referencias temporales;
- ejecutar un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 14.
- 35 16. Un producto de programa informático, almacenado en un medio no transitorio legible por ordenador, para detectar anomalías en un sistema de distribución de agua compuesto por una red de nodos, comprendiendo dicho producto de programa informático instrucciones de código para ejecutar un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 14.



TÉCNICA ANTERIOR

FIG.1

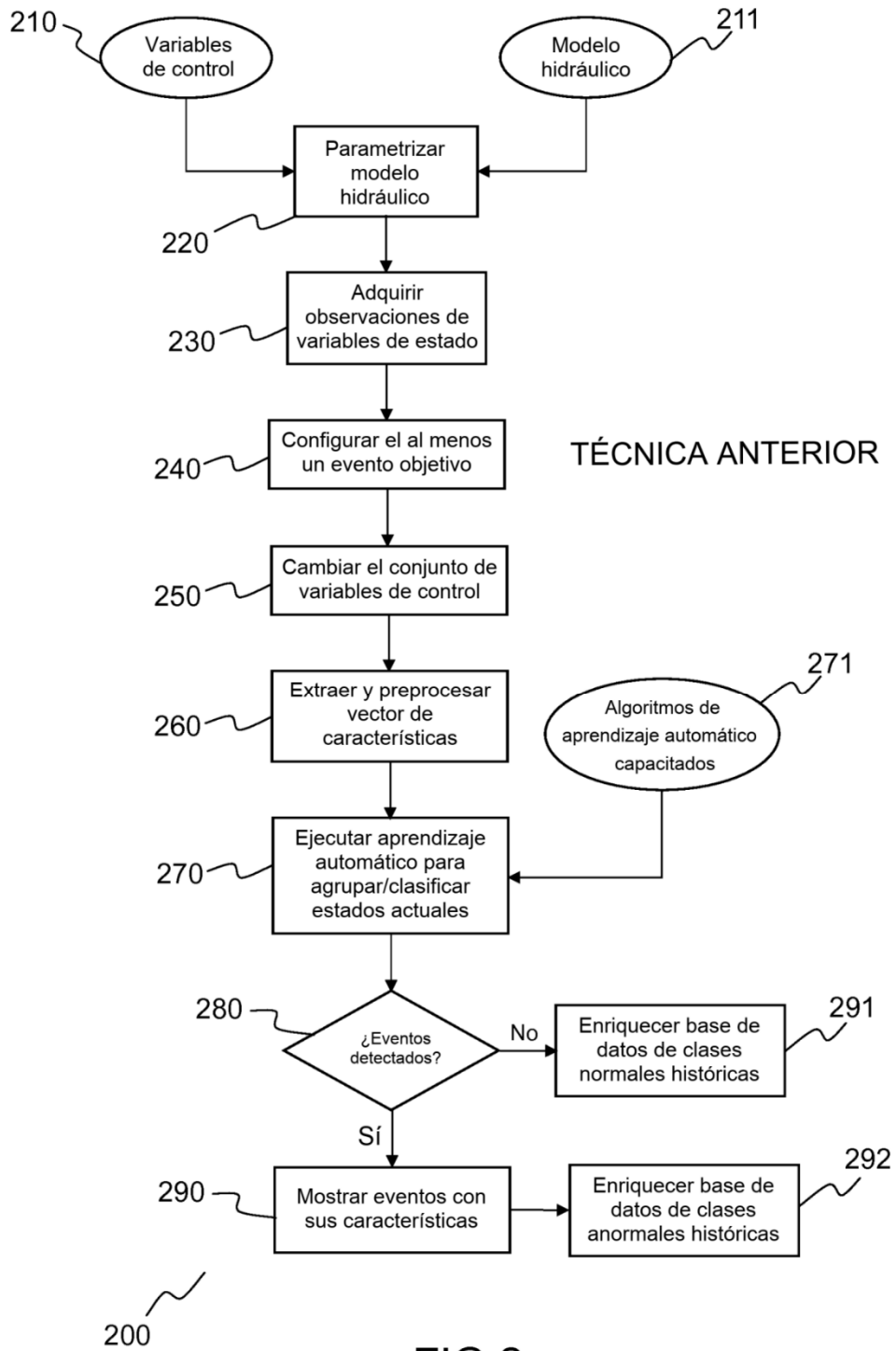
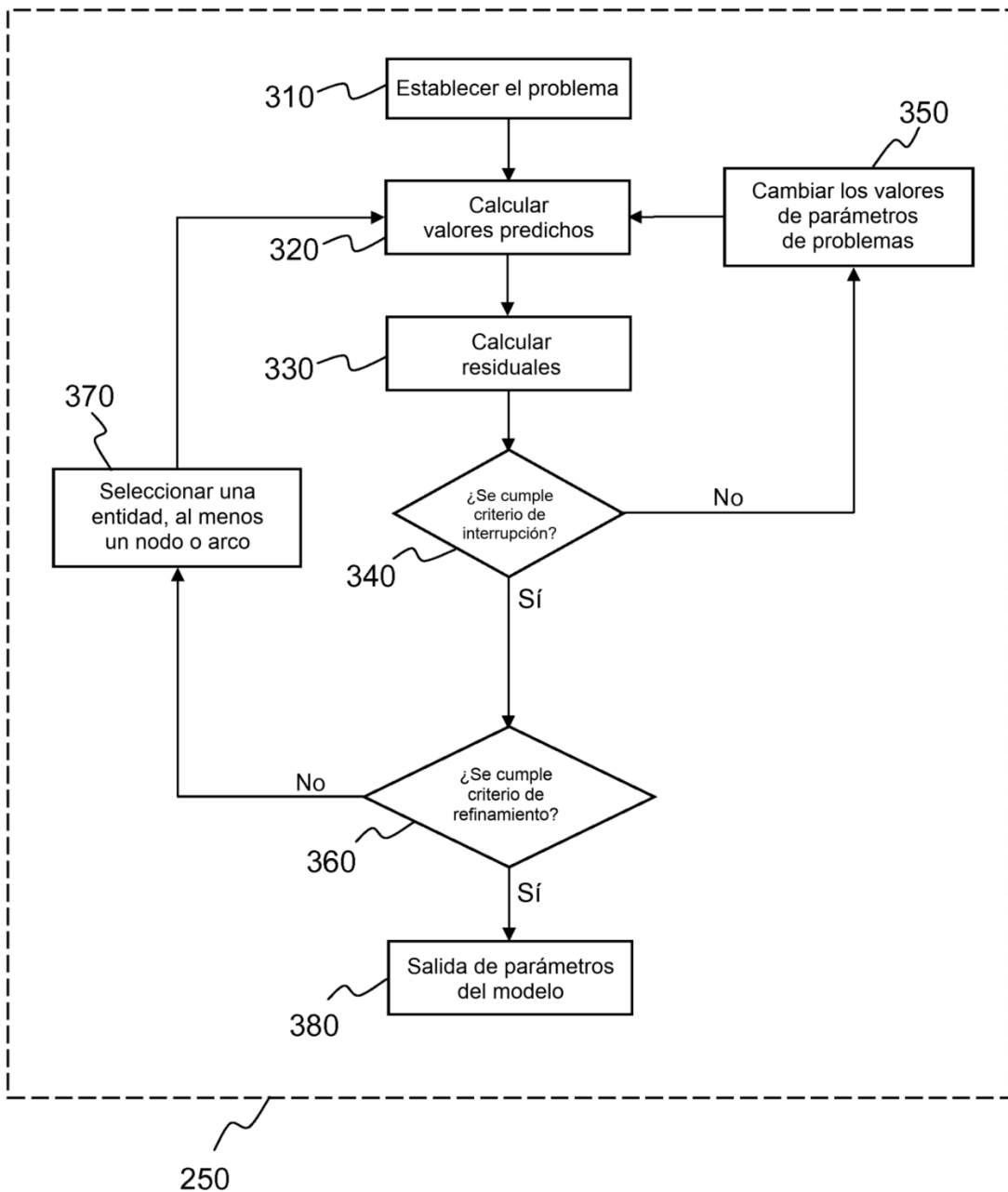


FIG.2



TÉCNICA ANTERIOR

FIG.3

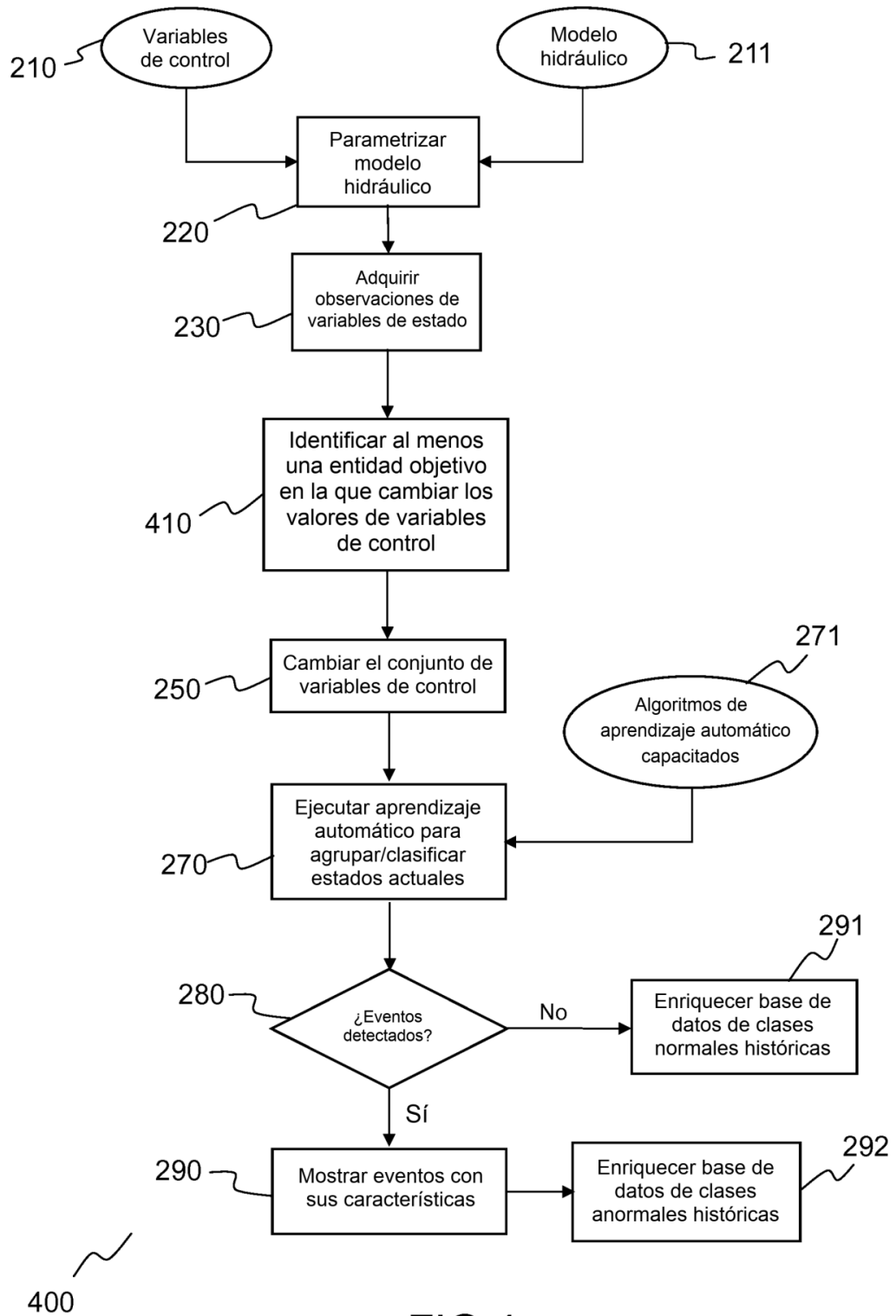


FIG.4

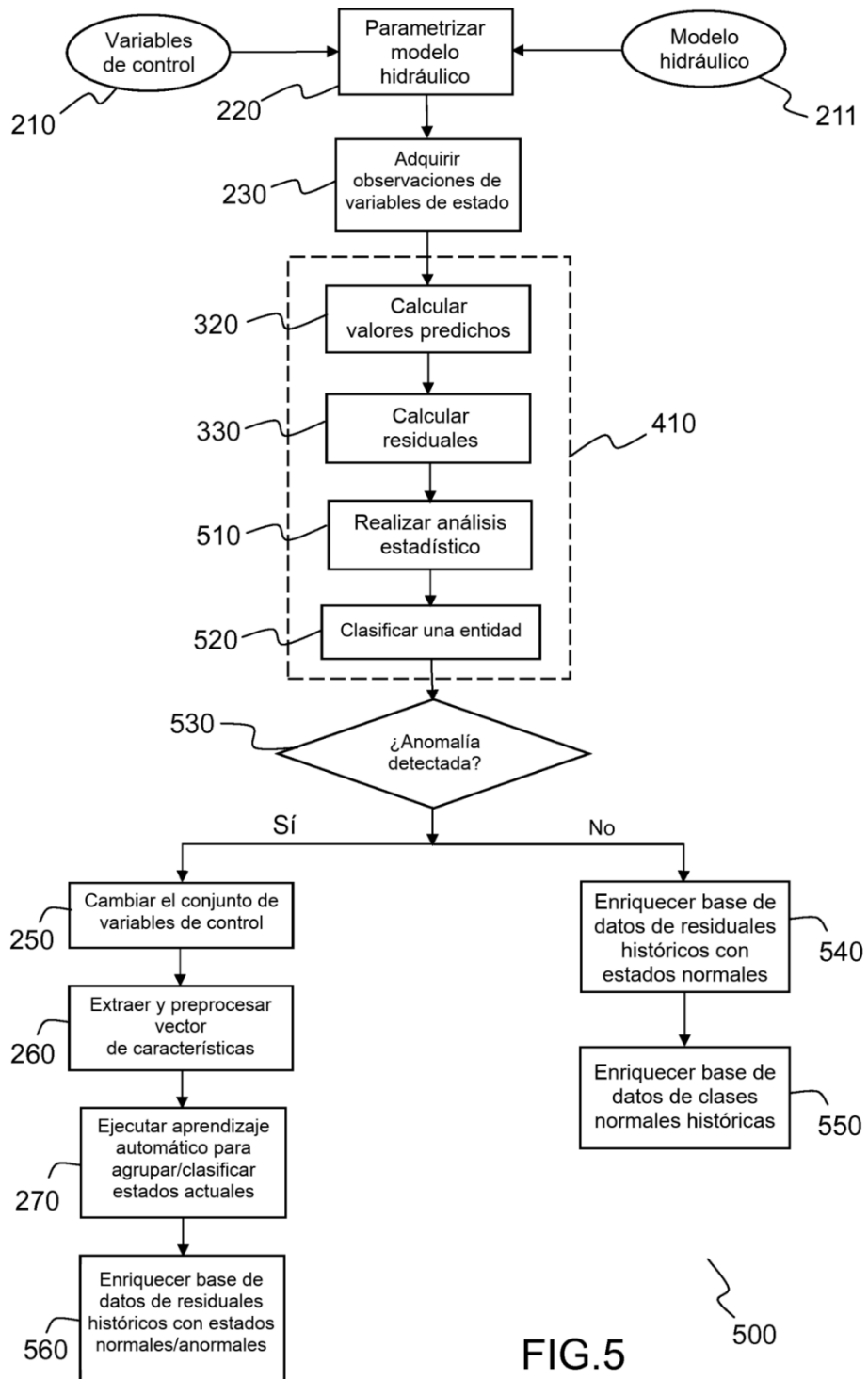


FIG.5

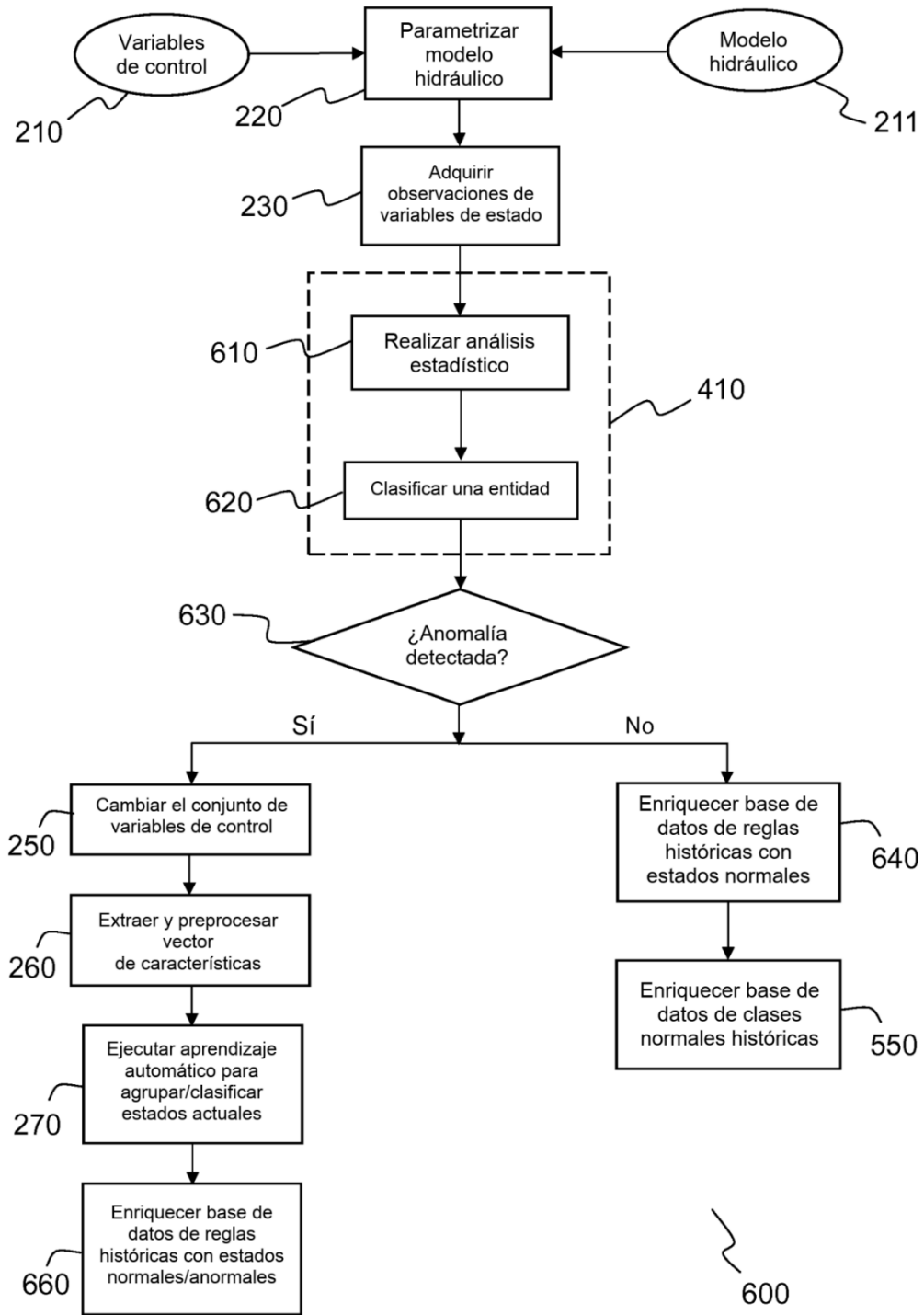


FIG.6

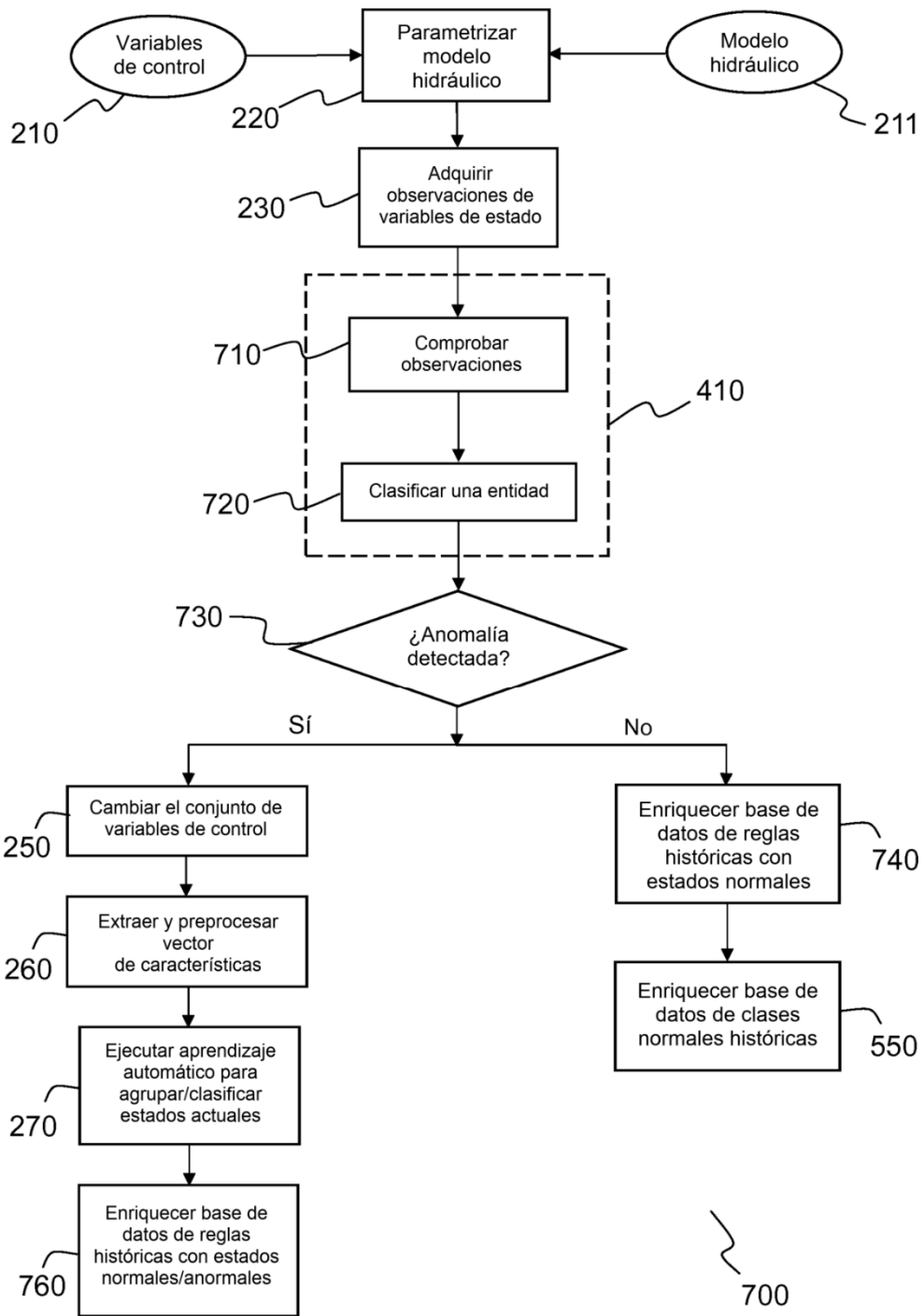


FIG.7

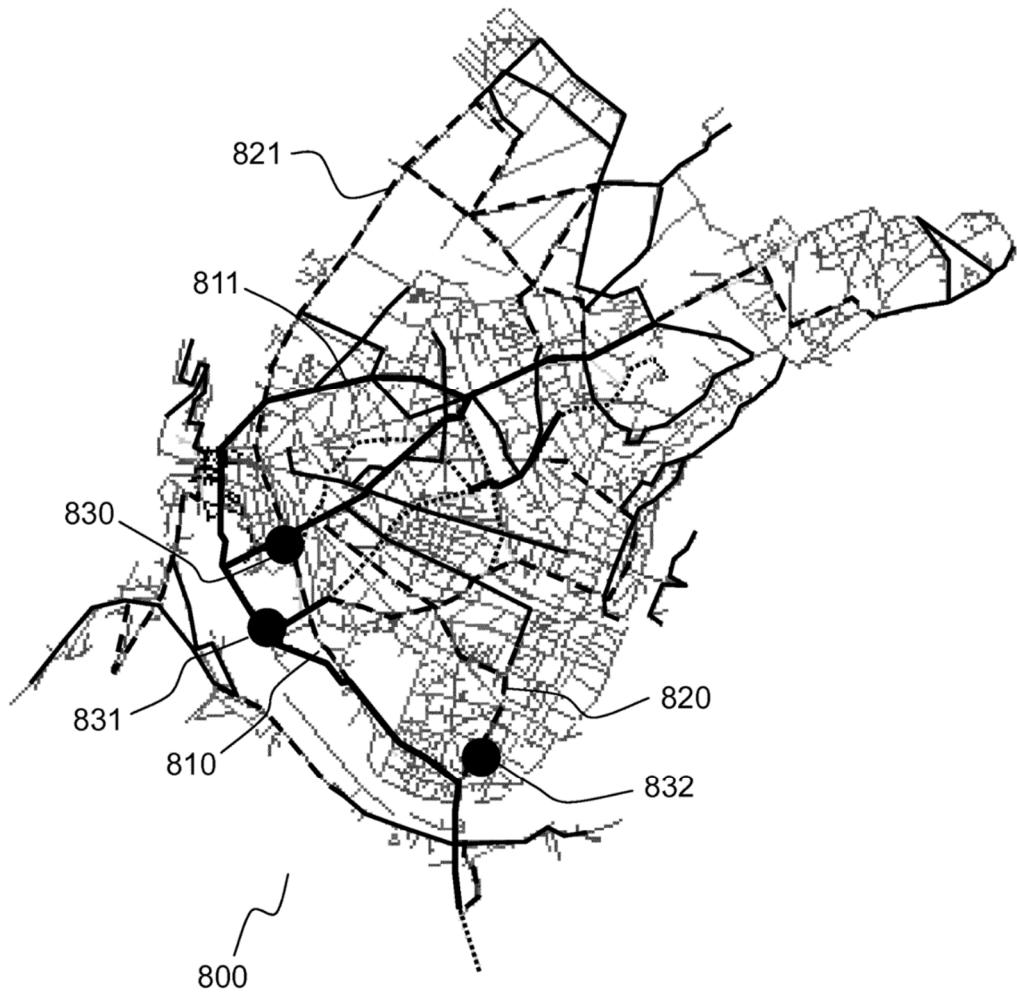


FIG.8